USING HIGH-THROUGHPUT GENE SURVEYS TO INVESTIGATE THE STABILITY OF ANAEROBIC DIGESTION AND FERMENTATION PROCESSES

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Catherine May Spirito

August 2017

USING HIGH-THROUGHPUT BIOINFORMATICS TO INVESTIGATE THE STABILITY OF ANAEROBIC DIGESTION AND FERMENTATION PROCESSES

Catherine May Spirito, Ph. D.

Cornell University 2017

In the carboxylate platform, undefined mixed cultures of anaerobic microbes can be used to convert organic wastes to fuels or valuable industrial chemicals. Anaerobic digestion is an established carboxylate platform technology in which a diverse consortia of microbes break down organic wastes to ultimately produce the gaseous end-product methane. More recently under the carboxylate platform umbrella, researchers have examined directing the anaerobic reactor microbiome toward producing medium-chain carboxylic acids instead of methane, by lowering the reactor pH and utilizing product extraction. The process by which these carboxylic acids are produced is called biological chain elongation. These medium-chain carboxylic acids, such as *n*-caproic and *n*-caprylic acid, can be used as antimicrobials in animal feed, in fragrances and flavors, or potentially be upgraded to diesel and aviation fuel. In this dissertation, I will present three studies which looked at the stability of carboxylate platform processes. In the first study, the effect of the dairy antibiotic monensin on the anaerobic digestion of cow manure was investigated. Monensin altered the anaerobic digestion microbiome and decreased methane production, though the digesters were able to adapt during the operating period. In the second study, the effect of changing substrate ratios of ethanol to acetate on the production of medium-chain carboxylic acids and the anaerobic reactor microbiome was investigated. The study found that higher substrate ratios of ethanol to acetate led to increased selectivity for *n*-caprylic acid production. In the third study, differences in the microbial communities of anaerobic fermenters that were producing medium-chain carboxylates from ethanol and acetate were investigated using Illumina 16S rRNA

gene sequencing, shotgun metagenomics, and metaproteomics. In both the second and third studies, I observed that shaping the anaerobic reactor microbiome for medium-chain carboxylate production led to a relatively uneven community, which was dominated by relatively few microbes. This work points toward the need to further improve the stability of the biological chain-elongation process.

BIOGRAPHICAL SKETCH

Catherine Spirito grew up in southern Rhode Island. She earned a BEng in Civil Engineering from McGill University in Montreal, Quebec in 2010. She began her PhD studies in Biological and Environmental Engineering at Cornell University in 2011. During her study, she worked as a research assistant in the lab of Dr. Lars Angenent. She received an EPA STAR Graduate Fellowship that funded two years of her studies (2014 to 2016). For two semesters, she served as a teaching assistant in the Department of Biological and Environmental Engineering, one semester for a lower-level undergraduate bioengineering thermodynamics and kinetics course and one semester for an upper-level undergraduate sustainable bioenergy systems capstone design course. She also served for two semesters as a teaching assistant for an introduction to comparative anatomy and physiology auto-tutorial course in the Department of Biology. For two years at Cornell, she shared her teaching experience with her fellow graduate students as a Graduate Teaching Assistant Fellow with the Center for Teaching Excellence.

I dedicate this dissertation to my partner, Pete Monacelli, and my family – my parents, Maureen Sullivan and John Spirito and my sister, Rachel Spirito. I greatly appreciate all their support and encouragement over the years.

ACKNOWLEDGMENTS

I thank my advisor, Lars Angenent, for encouraging and supporting me throughout my PhD. I really appreciate the lab environment he created. I really enjoyed working with all the people he recruited to work in our lab from across the country and around the world. I also thank my special committee members, James Gossett and Jeffrey Werner, for their support. I greatly enjoyed the biological processes course by James Gossett, which I took early on in my PhD at Cornell. Jeffrey Werner has provided me with invaluable support as I learned bioinformatics. I also appreciate him sharing his insights on undergraduate teaching. I also would like to thank all the members of the Angenent Lab, past and present. Especially, I would like to thank Joseph Usack for patiently teaching me everything he knew about running reactors. I would like to thank Michaela Teravest and her husband Andrew, as well as Matt Agler and Miriam Rosenbaum, for being so welcoming when I first moved to Ithaca. I would like to thank everyone in the lab who contributed countless hours helping me to design and build the stainless-steel reactors: Juan Guzman, Chase Brett, and Doug Caveney. I would like to thank the undergraduates and visiting scholars who worked with me and contributed greatly to my research: Sarah Daly, Alex Marzilli, and Leticia Regueiro. I would also like to thank Hanno Richter for all his thermodynamics help; Bastian Molitor and Mytien Nyugen for microbiology advice; Jiajie Xu for helping to take care of my reactors when I went on vacation; Leo Kucek for always being so enthusiastic about science and a great friend; and Lauren Harroff for always fixing the GCs when they broke, for being an awesome colleague and friend, and for providing me with so much support over these last few years.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION: CENTRAL AIM AND SUMMARY OF RESEARCH

1.1 Central Aim

In the context of the carboxylate platform, organic wastes can be converted to useful products by anaerobic reactor microbiomes (*i.e.*, open cultures of microbial consortia). Anaerobic digestion is one example of an established carboxylate platform technology and involves a microbial food web that converts the carbon in organic wastes to gaseous methane (Speece 1996). Reverse beta-oxidation in which ethanol or another electron donor drives the chain elongation of short-chain carboxylates to medium-chain carboxylates, such as *n*-caproate and *n*-caprylate, is an example of a promising carboxylate platform technology that is currently in development (Angenent et al. 2016, Grootscholten et al. 2013b, Kucek et al. 2016c). From an operating standpoint, it is desirable to know what contributes to a stable system for both these technologies. Therefore, the central aim of this research was to apply meta-omics tools, such as high throughput 16S rRNA gene sequencing, shotgun metagenomics, and metaproteomics, to conduct time-series studies on the effect of perturbations and imposed operating parameters on the anaerobic reactor microbiomes and system performance. Specifically, I looked at the effect of the dairy antibiotic monensin on anaerobic digestion of dairy cow manure. I also examined the effect of different ethanol to acetate substrate ratios on *n*-caprylate production in an upflow anaerobic filter reactor with inline product extraction. Finally, I examined the effect of different hydrogen partial pressures on the performance of these biological chain-elongation bioreactors, as well as on their microbial community structure.

1.2 Introduction

This dissertation describes previous work on the carboxylate platform and new work performed to further the understanding of the microbiomes involved. Chapter 2 is a literature review introducing previous work on anaerobic digestion and medium-chain carboxylate production, with a focus on thermodynamics and the use of microbial ecology tools to characterize the microbiome. Chapters 3, 4, and 5 describe the three aims of my dissertation work, and experiments performed toward those aims. Chapter 3 describes experiments performed with 5-L laboratory-scale anaerobic digesters to examine the effect of the dairy antibiotic monensin on the anaerobic digestion microbiome. Chapter 4 describes experiments performed with a 0.7-L laboratory-scale anaerobic fermenter with in-line product extraction to examine the effect of different ethanol to acetate ratios in the substrate on medium-chain carboxylate production. Chapter 5 describes experiments with 5-L laboratory-scale anaerobic fermenters with in-line product extraction to examine the impact of hydrogen partial pressures on process performance, as well as to further examine the microbial community involved in *n*-caprylate production. In Chapters 3, 4, and 5, Illumina 16S rRNA gene sequencing was used to characterize the microbiome. In Chapter 5, shotgun metagenomics and metaproteomics were also employed. Appendix 1, 2, and 3 contains supplementary material for Chapters 3, 4, and 5, respectively. Appendix 4 includes experiment protocols: a protocol for preparing environmental samples for Illumina 16S rRNA gene sequencing; and a protocol for initial analysis of the resulting sequences.

1.3 Summary of Experiments

Section 1-Aim: To examine the effect of the dairy antibiotic monensin on the anaerobic

digestion microbiome.

- 5-L continuously-stirred, anaerobic digesters fed with dairy cow manure with and without monensin were operated for a period of approximately one year.
- Standard physical and chemical parameters, such as alkalinity, biogas production, pH, total ammonium, and volatile fatty acid production were monitored.
- Illumina 16S rRNA gene sequencing was performed of biomass samples from the anaerobic digesters and from the cows that were dosed with monensin.

Section 2- Aim: To examine the effect of different ethanol to acetate substrate ratios on the production of medium-chain carboxylates in a bioreactor with product extraction.

- A 0.7-L upflow anaerobic filter bioreactor with in-line product extraction was fed with different ratios of ethanol and acetate in the substrate
- Medium-chain carboxylic acid production rates were monitored, as well as gas production, reactor pH, and temperature.
- Illumina 16S rRNA gene sequencing was performed of biomass samples collected from the bottom and middle ports of the bioreactor.

Section 3- Aim: To examine the impact of hydrogen partial pressures on the microbial community involved in the production of *n*-caprylate from ethanol and acetate.

- Three 5-L stainless-steel bioreactors with in-line product extraction were constructed.
- The bioreactors were operated for a period of approximately 9 months and medium-chain carboxylic acid production rates were monitored under different hydrogen partial pressures.
- Illumina 16S rRNA gene sequencing, shotgun metagenomics, and metagenomics were performed of biomass samples collected in a time-series from the bioreactors.

CHAPTER 2

CARBOXYLATE PLATFORM WITH ANAEROBIC REACTOR MICROBIOMES: A LITERATURE REVIEW

Portions of this literature review were adapted from: Spirito, C. M., H. Richter, et al. (2014). "Chain elongation in anaerobic reactor microbiomes to recover resources from waste." Current Opinion in Biotechnology **27**(0): 115-122.

2.1 Introduction

Biorefinery platforms are designed to take biomass and convert it into useful fuels, heat, power, or value-added chemicals. The sugar platform is one example in which enzymes convert biomass into five-to-six carbon sugars, which then can be converted further to fuels or other chemicals. Another is the syngas platform in which thermochemical processes are used to convert biomass to syngas (*i.e.*, a mixture of carbon dioxide, carbon monoxide, and hydrogen gas), which can then be directed toward fuel production. A third platform is the carboxylate platform in which open cultures of microbial consortia, which are referred to here as anaerobic reactor microbiomes, break down and convert organic waste materials to a mixture of carboxylates. Here, I will use the term carboxylates to include both the dissociated and undissociated form of carboxylic acids, which are organic acids with at least one carboxyl group. These carboxylates represent intermediate platform chemicals that can then be further upgraded to fuels or value-added chemicals. The carboxylate platform simultaneously addresses two pressing issues of our time: declining fossil fuel reserves and the need for proper waste management. In addition, there are many advantages to the use of

the carboxylate platform in a biorefinery context. The carboxylate platform utilizes an open and anaerobic microbiome, which circumvents the costly need for sterilization or aeration. In addition, the diverse communities that are present in anaerobic bioreactors of the carboxylate platform are able to handle various complex organic waste streams (Agler et al. 2011, Holtzapple and Granda 2009).

Anaerobic digestion is one established example of a carboxylate platform process. In the context of anaerobic digestion, a complex microbial food web breaks down organic substrates *via* hydrolysis, acidogenesis, acetogenesis, and methanogenesis to produce the main end product gaseous methane, which easily bubbles out of solution. Methane can be used for electrical power generation or upgraded to a higher quality to inject into natural gas pipelines. Though anaerobic digestion is a proven process, the low price of electricity in recent years and the high capital investment to build a system in the US has made anaerobic digestion less economically attractive (Agler et al. 2014). In their economic policy analysis paper on farm-based anaerobic digestion in the US, Zaks et al. (2011) found the net present value of anaerobic digester systems is too low for widespread adoption. The authors did note that a climate policy that put an economic value on environmental and energy benefits of a technology could help to increase adoption.

The carboxylate platform also includes processes broadly termed biological chain elongation by which carbon atoms are added to shorter chain carboxylates and other carbon compounds. For example, homoacetogenesis is a chain-elongation process in which two carbon dioxide molecules are combined to form acetate. Reverse beta-oxidation is another chain-elongation process for which ethanol and potentially other electron donors (*e.g.*, lactate (Spirito et al. 2014)) can be used to upgrade acetate and other short-chain carboxylates derived from organic waste materials to longer chain carboxylates such as *n*-caproate. This chain elongation is beneficial

since *n*-caproate is more energy dense and hydrophobic (*i.e.*, easier to separate from solution) than the precursor short-chain carboxylates. *n*-Caproate can also be further upgraded to liquid fuels (*e.g., via* ketonization to *n*-undecane), industrial solvents, or other bulk chemicals. It can be used as animal feed, a green antimicrobial, or as an anticorrosive agent (Agler et al. 2012a, Agler et al. 2011, Butkus et al. 2011, Gonzalez-Cabaleiro et al. 2013, Holtzapple and Granda 2009, Levy et al. 1981, Marshall et al. 2013). Recent work has demonstrated that the effluent from syngas fermentation, which contains dilute ethanol and acetate, can be used as substrate for the reverse beta-oxidation process (Vasudevan et al. 2014). By converting dilute ethanol (typically around 2-4% w/w for pure cultures; (Richter et al. 2013)) from syngas fermentation to *n*-caproate, which is a relatively hydrophobic chemical, energy intensive ethanol distillation can be avoided.

Though *n*-caproate production has been demonstrated successfully in lab-scale systems with production rates comparable to solid state anaerobic digestion (Agler et al. 2012a, Steinbusch et al. 2011), there is still room for improvement before it is scaled up. One area for improvement is lowering the electron donor (*e.g.*, ethanol) to product (*e.g.*, *n*-caproate) ratio. As will be covered later in this review, a better understanding of the effect of hydrogen partial pressure on the system may enable researchers to more successfully direct the *n*-caproate production pathway and lower the electron donor to product ratio further.

In recent years, there has been an increasing recognition of the need to no longer look at reactor systems as black boxes and to examine the correlations between changes in the reactor microbiome and changes in reactor performance (Koch et al. 2014). In previous work in the Angenent lab, researchers statistically linked operating and environmental parameters with the performance and microbiome in anaerobic digesters (Werner et al. 2011) and *n*-butyrate producing reactors (Agler et al. 2012b). As will be discussed later, work still needs to be undertaken to

understand what leads to a stable and functionally redundant system in medium-chain carboxylate producing bioreactors and other chain-elongation processes. In shaping these systems to perform a specific function, it is important to ensure that the stability of the system is not compromised.

This review will broadly cover the effect of perturbations on the reactor microbiome and performance of the carboxylate platform system. A look at three carboxylate platform pathways: anaerobic digestion, homoacetogenesis, and reverse beta-oxidation will be provided. This review will cover the metabolism, thermodynamics, and research and industrial applications of these pathways, with an emphasis on reverse beta-oxidation. A brief overview of the effect of monensin on anaerobic digestion will be given. A focus will be placed on the effect of hydrogen partial pressures on reactor microbiomes carrying out reverse beta-oxidation. Finally, the use of microbial ecology tools to explore the concept of reactor stability and the effect of disturbances on the microbiome and system performance will be discussed.

2.2 Carboxylate Platform Pathways

2.2.1 General Thermodynamics and Metabolism

In anaerobic fermentations, bacteria often operate at the minimum energy level required to carry out biological reactions. Due to competition, microbes are forced to maximize their thermodynamic efficiency of ATP synthesis (Buckel and Thauer 2013). The synthesis of ATP requires about 60 to 70 $kJ \cdot mol^{-1}$, if heat losses are considered. Generally, the minimum amount of energy a living cell can make use of is equivalent to the transfer of one ion across the cytoplasmic membrane, or one-third of an ATP unit. This works out to a minimum energy production requirement of exergonic reactions in cells of ~-20 $kJ \cdot mol^{-1}$ (Schink 1997). In a study on mixed cultures carrying out reductive dechlorination (Fennell and Gossett 1998), the authors found that

organisms in these cultures required a minimum negative free energy of -19 $kJ \cdot mol^{-1}$ to live and grow. More recent work indicates that some pathways may remain feasible at ΔG_r values greater than -10 $kJ \cdot mol^{-1}$, including *n*-butyrate oxidation (Jackson and McInerney 2002, Kleerebezem and Stams 2000).

Previously, anaerobic fermentations were thought to only generate ATP *via* substrate level phosphorylation (SLP) in which ATP is formed *via* transfer of a phosphoryl group from a substrate to ADP. However, a third type of energy conservation, which is termed flavin based electron bifurcation (FBEB) due to its complex of four flavin adenine dinucleotides (FAD), has recently been discovered (Herrmann et al. 2008). This form of energy conservation is important in the pathways discussed in this paper: hydrogenotrophic methanogenesis; homoacetogenesis; and reverse beta-oxidation. FBEB will be discussed in more detail here in regards to reverse beta-oxidation. Briefly, FBEB (**Figure 2.1**) involves the coupling of the endergonic reduction of low potential ferredoxin *via* higher potential electron donors, such as NAD(P)H or H_2, with the exergonic reduction of higher potential electron acceptors (*i.e.*, crotonyl-CoA, $NAD^{\pm}$, or heterodisulfide) *via* the same electron donor (*i.e.*, NAD(P)H or H_2). This reduced ferredoxin can then be reoxidized in several ways. Certain bacteria, including *Clostridium kluyveri*, which contain the membrane associated Rnf complex, can reoxidize the ferredoxin *via* $NAD^{\pm}$, which drives proton or sodium ion translocation and leads to the synthesis of ATP (Buckel and Thauer 2013).

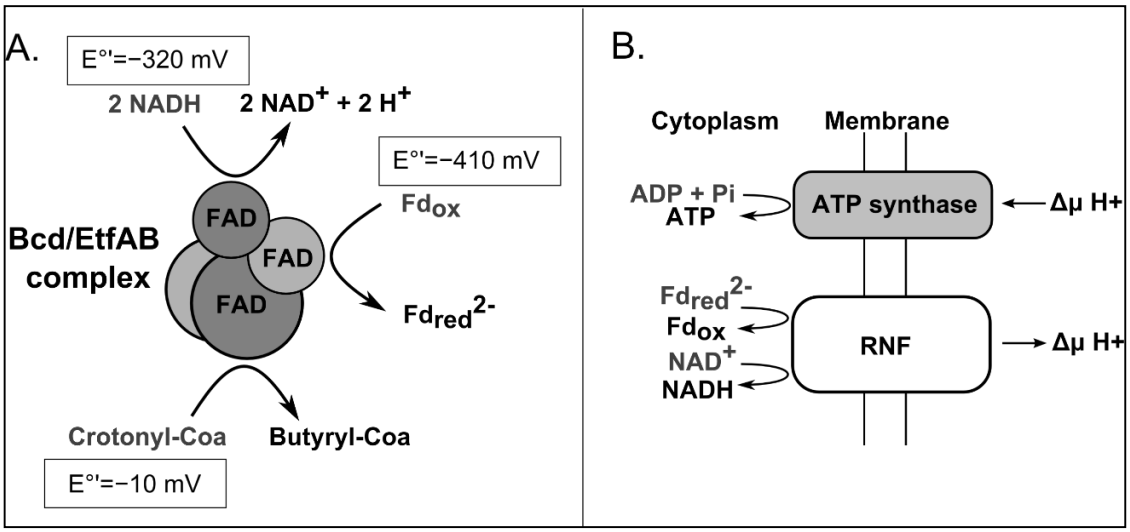

Figure 2.1. The complex (Bcd/EtfAB) responsible for FBEB in C. kluyveri (**A**) and the RNF complex with ATP synthase (**B**). Figure adapted from Li et al. (2008) and Biegel et al. (2011). Further explanation of the Figure can be found in Section 2.5.1.

The feasibility of reactions carried out in anaerobic fermentations is governed by the Gibbs free energy of the reaction (ΔG_r) since most of these reactions operate close to thermodynamic equilibrium (Kleerebezem and Van Loosdrecht 2010). The standard Gibbs energy change for formation of a compound ($\Delta G°_f$) at temperature (T) can be expressed as:

$$\Delta G_f^\circ = \Delta H_f^\circ - T\Delta S_f^\circ$$

where $\Delta S°_f$ represents the standard entropy of formation and $\Delta H°_f$ the standard enthalpy of formation of the compound at that temperature. Gibbs free energy of formation and enthalpy of formation values used in this review were found in Kleerebezem and Loosdrecht (2010) and the CRC handbook for Chemistry and Physics, 94th *ed.* (2013). Under standard conditions (25°C, 1 atm CO_2 and H_2, 1 M all other reactants and products), the standard Gibbs free energy of a reaction ($\Delta G_r°$) can be calculated as:

$$\Delta G_r^\circ = \left(\sum_{i=1}^{n} v_i * \Delta G_{fi}^\circ\right)_{products} - \left(\sum_{i=1}^{n} v_i * \Delta G_{fi}^\circ\right)_{reactants}$$

where v_i represents the stoichiometric coefficient and $\Delta G°_{fi}$ represents the free energy of formation for the reaction constituent i. Likewise, the standard enthalpy of a reaction ($\Delta H_r°$) can be calculated as:

$$\Delta H_r^\circ = \left(\sum_{i=1}^{n} v_i * \Delta H_{fi}^\circ\right)_{products} - \left(\sum_{i=1}^{n} v_i * \Delta H_{fi}^\circ\right)_{reactants}$$

The Gibbs-Helmholtz equation (Kleerebezem and Van Loosdrecht 2010) can be used to calculate the $\Delta G_r°$ at a specified temperature (T2) based on the calculated standard values for $\Delta G_{T1}°$ and $\Delta H_{T1}°$ at 25°C where:

$$\Delta G_{T2}^\circ = \Delta G_{T1}^\circ \left(\frac{T2}{T1}\right) + \Delta H_{T1}^\circ \left(\frac{T1 - T2}{T1}\right)$$

As mentioned above, the standard Gibbs free energy of a reaction ($\Delta G_r°$) assumes that the concentrations of the reactants and products are at 1 M, or 1 atm in the case of gases. Therefore, to gain an accurate insight into the feasibility of a reaction it is necessary to correct for the actual activities of reactants and products in a system when calculating the Gibbs free energy of the reaction (ΔG_r):

$$\Delta G_r = \Delta G_r^\circ + RTlnQ$$

where R is the universal gas constant (8.3145×10^{-3} kJ·K^{-1}·mol^{-1}), T represents the temperature of the solution in degrees Kelvin, and Q is the ratio of activities of the products to reactants at equilibrium. The activities of the constituents are dependent on the ionic strength of the reactor solution (Kleerebezem and Van Loosdrecht 2010). Replacing Q and using a_i to represent the activity of the constituent i in solution and v_i to represent the stoichiometric coefficient of the

constituent i (where v_i is negative for reactants and positive for products), the equation can be written as:

$$\Delta G_r = \Delta G_r^\circ + RTln\left(\prod a_i^{v_i}\right)$$

As mentioned previously, this ΔG_r must be less than or equal to approximately -20 kJ·mol^{-1} for this reaction to meet the minimum energy production requirement of the cell. For anaerobic fermentation systems, this equation can be used to predict the different hydrogen partial pressures below which *n*-butyrate or ethanol oxidation is feasible (CRC 2013, Hanselmann 1991, Kleerebezem and Van Loosdrecht 2010).

For biological systems, the standard Gibbs free energy of a reaction can be corrected for pH 7 and referred to as $\Delta G_r^{\circ\prime}$, where:

$$\Delta G_r^{\circ\prime} = \Delta G_r^\circ + RTln\{H^+\}^q$$

where $\{H^+\}$ is the activity of the hydrogen ions in solution, and q is the number of hydrogen ions exchanged in the reaction, with a positive value if they are produced and negative value if they are consumed in the reaction. This equation can be further simplified to:

$$\Delta G_r^{\circ\prime} = \Delta G_r^\circ + qRTln(10^{-pH})$$

where pH is the pH of the aqueous solution (*i.e.,* pH 7). The values of $\Delta G_r^{\circ\prime}$ for reactions of interest to this literature review are tabulated in **Table 2.1**.

Table 2.1. Thermodynamic information for oxidation reactions (carboxylates and ethanol), homoacetogenesis, and reverse beta-oxidation, which may play a role in anaerobic fermentation systems. All $\Delta G_r^{\circ'}$ values are calculated considering all reactants and products to be in the aqueous phase except for H_2 and CO_2, which are gaseous at 101.33 kPa. $\Delta G_r^{\circ'}$ values are at biological standard state (pH=7 at 25°C) and calculated using ΔG_f° and ΔH_f° values for the individual compounds at 25°C from Kleerebezem and van Loosdrecht (2010) and CRC Handbook of Chemistry and Physics, 94th ed. (2013), and the Gibbs-Helmholtz equation. For chain-elongation reactions with ethanol oxidation, the reactions are shown with their individual and overall $\Delta G_r^{\circ'}$ values. Two equations (*i.e.*, a and b) for reverse beta-oxidation to *n*-butyrate are included here (Buckel and Thauer 2013). Additional information on these two reactions can be found in Section 5: Reverse Beta-Oxidation.

Pathway	**Reactions**	**Coupled Reactions**	**$\Delta G_r^{\circ'}$ (kJ·mol^{-1})**
n-Butyrate oxidation	*n*-butyrate$^-$ + 2 H_2O → 2 acetate$^-$ + H^+ + 2 H_2	-	48.24
n-Caproate oxidation	*n*-caproate$^-$ + 2 H_2O → *n*-butyrate$^-$ + acetate + 2 H_2 + H^+	-	48.44
Homoacetogenesis	4 H_2 + 2 CO_2 → acetate$^-$ + H^+ + 2 H_2O	-	-94.96
Ethanol oxidation	ethanol + H_20 → acetate$^-$ + H^+ + 2 H_2	× 1	9.64
Reverse beta-oxidation	ethanol + acetate$^-$ → *n*-butyrate$^-$ + H_2O	× 5	-193.00[1]
Chain elongation of acetate and ethanol to *n*-butyrate (a)	6 ethanol + 4 acetate$^-$ → 5 *n*-butyrate$^-$ + H^+ + 2 H_2 + 4 H_2O	Overall	-183.36[1]
Chain elongation of acetate and ethanol to *n*-butyrate (b)	5 ethanol + 3 acetate$^-$ → 4 *n*-butyrate$^-$ + 1 H^+ + 3 H_2O + 2 H_2	Overall	-144.76[2]
Ethanol oxidation	ethanol + H_20 → acetate$^-$ + H^+ + 2 H_2	× 1	9.64
Reverse beta-oxidation	ethanol + *n*-butyrate$^-$ → *n*-caproate$^-$ + H_2O	× 5	-194.00[1]
Chain elongation of *n*-butyrate and ethanol to *n*-caproate	6 ethanol + 5 *n*-butyrate$^-$ → acetate$^-$ + 5 *n*-caproate$^-$ + H^+ + 2 H_2 + 4 H_2O	Overall	-184.36[1]

[1] The unit is kJ/6 mol of ethanol; [2] The unit is kJ/5 mol of ethanol

Another important consideration in carboxylate platform systems is the pH of the reactor liquid since carboxylates are weak organic acids, which appear more in their undissociated or dissociated form based on pH, as determined by the Hendersson-Hasselbalch equation. For example, *n*-caproate has a pKa value of 4.85 at 25°C (CRC 2013). At lower pH values, there is a

shift toward more of the undissociated form in solution. The undissociated form of the acid can cross the microbial cell membrane and dissociate inside the cell, due to the more alkaline pH inside the cytoplasm. The increase in intracellular anion concentrations can have inhibitory effects on the microbial cell, which must work to try to keep a constant proton gradient across its membrane (Palmqvist and Hahn-Hägerdal 2000). If continuous bioreactors are run at low pH values, it is often necessary to extract these carboxylic acids out of solution to prevent end product toxicity (Agler et al. 2012a).

2.2.2 General Kinetics

For a reaction to occur in a biological system, thermodynamic limitations (*i.e.*, $\Delta G_r \leq \sim$ -20 kJ·mol^{-1}) and kinetic limitations (*i.e.*, there must be a high enough concentration of enzyme catalyst present to drive the reaction) must be satisfied. Ideally, both these limitations should be considered in fermentation models. Therefore, here I will briefly discuss fermentation kinetics. Early studies on anaerobic fermentation often turned to Monod kinetics to describe the kinetics of the reactions involved (Rodríguez et al. 2008). The Monod equation states that:

$$\mu = \mu^{max} * \frac{S}{K_s + S}$$

Where μ (h^{-1}) is the specific growth rate, μ^{max} (h^{-1}) is the maximum specific growth rate, S (mol·L^{-1}) is the concentration of limiting substrate, and K_s (mol·L^{-1}) is the half saturation constant for microbial growth (Kleerebezem and Stams 2000). Kinetic models based on Monod kinetics have been successfully applied in aerobic systems and other systems operating far from thermodynamic equilibrium. In these systems, due to the large (negative) standard $\Delta G_r°$ values, the driving force of catabolic reactions will remain high despite potential changes in substrate and product concentrations. The substrate conversion and growth can be considered coupled, and Monod

kinetics can be applied. However, for systems operating close to thermodynamic equilibrium, such as in anaerobic fermentation systems, the substrate and product concentrations can have a considerable impact on the driving force of the reaction. Often, the products can build up to inhibitory levels for the forward direction of the reaction (positive ΔG_r), which leads to the reverse reaction occurring. Therefore, substrate conversion and growth should be uncoupled in models, and Monod kinetics does not accurately describe the situation (Kleerebezem and Stams 2000). Increasingly, authors are incorporating both bioenergetics and kinetics into mixed culture fermentation models (Rodríguez et al. 2008).

2.2.3 Gas-Liquid Transfer Relations

Henry's Law governs the distribution of a chemical species between liquid and gas phase at equilibrium:

$$A_{dissolved} = H * p_A$$

Where $A_{dissolved}$ ($mol \cdot L^{-1}$) is the concentration of species A dissolved in the liquid, p_A (Pa) is the partial pressure of A in the gas phase, and H ($mol \cdot L^{-1} \cdot Pa^{-1}$) is the Henry's Law constant for that compound (Jones and Greenfield 1982). Henry's Law constants for relevant compounds are provided in **Table 2.2** for the gases of interest. As can be seen, hydrogen is much less soluble in water than carbon dioxide, with a Henry's Law constant of 7.4×10^{-9} $mol \cdot L^{-1} \cdot Pa^{-1}$ *vs.* 2.7×10^{-7} $mol \cdot L^{-1} \cdot Pa^{-1}$. Finally, carbon dioxide species distribution is also influenced by carbon dioxide-bicarbonate equilibrium in the reactor.

It is important to recognize that equilibrium relationships may not always govern in the bioreactors in question and gas-transfer limitations must be considered. As can be seen from **Table 2.2**, the diffusivity coefficient for hydrogen is relatively high. In bioreactor systems, several studies

(Beckers et al. 2015, Kraemer and Bagley 2006, Pauss et al. 1990) have reported differences between the theoretical liquid concentration of hydrogen calculated based on equilibrium with the headspace and the actual measured liquid concentration in these reactor systems. Changes in reactor design, such as improved gas headspace recirculation through the liquid phase, can be used to improve the overall volumetric mass transfer coefficient (k_La) for the hydrogen gas and overcome these gas-transfer limitations (Pauss et al. 1990).

Table 2.2. Diffusivity coefficients and Henry's constant values for gases involved in anaerobic fermentation at 25°C (Pauss et al. 1990).

Gas	**Diffusivity Coefficient (D) (10^5 $cm^2 \cdot s^{-1}$)**	**Henry's Constant (K_H) ($mol \cdot L^{-1} \cdot Pa^{-1}$)**
H_2	4.65	7.40×10^{-9}
CH_4	1.57	1.12×10^{-8}
CO_2	1.98	2.70×10^{-7}

2.3 Anaerobic Digestion

In anaerobic digestion, organic polymers, such as proteins, polysaccharides, and lipids, are hydrolyzed to monomers and oligomers such as amino acids, sugars, and glycerol. This process occurs *via* extracellular hydrolytic enzymes produced by primary fermentative bacteria. These primary fermenting bacteria then further break down the monomers and oligomers to low weight organic acids and alcohols, specifically fatty acids, succinate, lactate, alcohols, acetate, hydrogen, and carbon dioxide, in a process termed acidogenesis. Acetate, hydrogen, carbon dioxide, and other one-carbon compounds can be converted directly to carbon dioxide and methane by acetoclastic (*i.e.*, *via* acetate cleavage) and hydrogenotrophic (*i.e.*, hydrogen and carbon dioxide as substrates) methanogens. *Via* acetogenesis, the other longer-chain fatty acids and alcohols (termed intermediate products) can be further oxidized to hydrogen, carbon dioxide, formate, and

acetic acid, which are then converted *via* methanogenesis to methane and carbon dioxide by methanogens (Schink 1997). As methane is a gaseous product with a low solubility, it easily bubbles out of the bioreactor liquid, which prevents end product inhibition. Acetoclastic methanogens contribute to about two thirds of the methane formation in a typical anaerobic digester (Speece 1996).

Interspecies hydrogen transfer carried out by syntrophic communities of bacteria and archaea plays an important role in keeping the flux of the pathway toward producing methane. Hydrogenotrophic methanogens help to maintain low hydrogen partial pressures by consuming hydrogen, which in turn allows for carboxylate oxidation by fermenting bacteria and keeps the pathway directed toward methane formation. In the carboxylate oxidation reactions, the fermenting bacteria use protons as terminal electron acceptors and NADH as a redox mediator, which leads to the formation of hydrogen. Under standard conditions, it is thermodynamically unfavorable to reduce protons using a redox mediator such as NADH. However, in methanogenic environments, low hydrogen partial pressures make the oxidation of NADH coupled to proton reduction exergonic, which enables carboxylate oxidation. If hydrogen partial pressure is kept low in the system, the maximum substrate flux goes through acetate and methane because more ATP can be generated *via* that pathway than in the formation of intermediate products such as *n*-butyrate or propionate (Stams and Plugge 2009). Frequently, practitioners refer to a ‘stuck’ anaerobic digesters in which an upset, such as substrate overloading, leads to an increase in hydrogen partial pressure, which in turn causes increased formation of the intermediate products. Methanogens are not able to directly utilize these intermediate products, which leads to buildup of the products and acidification of the reactor. This results in a lowering of the reactor pH, which leads to inhibition of the methanogens (Harper and Pohland 1986, Schink 1997).

2.3.1 Effect of the Antibiotic Monensin

Monensin is a monovalent carboxylic ionophore antibiotic produced *via* fermentation by *Streptomyces cinnamonensis* and given to dairy cows to increase their milk production efficiency or to treat coccidiosis, which is a parasitic infection (Simjee et al. 2012). Monensin acts by altering fermentation pathways in the cow rumen. Specifically, monensin, similar to other ionophore antibiotics, inserts itself in the cell membrane of some bacteria, which interferes with the ion gradients needed to transport nutrients and generate proton motive force. In response, the bacterium diverts its ATP away from other essential cell processes to try to restore the normal ion gradient and growth ceases (Simjee et al. 2012). Generally, Gram positive bacteria are more sensitive to monensin than Gram negative bacteria, due to differences in the cell membrane. Previous studies have shown that monensin inhibits the formation of the precursors of methanogenesis in the cow rumen. Monensin increases propionate-producing bacteria and decreases hydrogen, formate, and acetate-producing bacteria, which can lead to decreased methane production in the cow rumen (Chen and Wolin 1979) (Simjee et al. 2012). For example, monensin has been found to select in favor of the propionate-forming *Selenomonas* spp. (Thaveesri et al. 1994).

A portion of the monensin is excreted in the cow manure, and this manure is fed to the farm-based anaerobic digester system. Previous studies and on-farm experiences have shown a decrease in methane production upon first introduction of monensin to anaerobic digesters (Thaveesri et al. 1994, Varel and Hashimoto 1981, Zitomer 2007). These studies indicated that monensin indirectly inhibits methanogenesis by depleting the precursor acetic acid. However, a six-month study by Varel and Hashimoto (1982) indicated the capability of anaerobic digesters

fed monensin-laden manure to adapt to this ionophore over time when the hydraulic retention time was slowly shortened. Pure culture studies (Simjee et al. 2012) have also indicated that bacteria can develop monensin resistance through cell wall or glycocalyx thickening, which is extracellular polymeric material produced by some bacteria. However, a lack of long-term studies in this area has limited understanding on how to prevent possible digester failure or decreased performance. In addition, few studies exist linking the effect of antibiotics on the anaerobic digester microbiomes to changes in digester performance.

A previous study (Angenent et al. 2008) on the effect of tylosin, which is a macrolide antibiotic, on anaerobic digestion of swine waste found that tylosin was degraded rapidly by an acclimated digester microbiome. The authors hypothesized that the fast degradation of the antibiotic was the main reason biogas production was not affected by the continual tylosin manure addition to the acclimated digester. There have been few previous reports on the rate of degradation of monensin in anaerobic digestion. A recent study (Varel et al. 2012) on the anaerobic digestion of cattle manure containing monensin concluded that anaerobic digestion at mesophilic temperatures (38°C) was not an effective method to reduce monensin concentrations. Over a 28-day period of running 2-L anaerobic digesters fed cattle manure containing 0.36 $mg \cdot L^{-1}$ monensin, they saw only an 8% decrease in the concentration of monensin exiting the anaerobic digester. However, this study did not investigate whether acclimation of the system to monensin would lead to an increased degradation rate. Studies on the half-life of monensin have mainly focused on monensin degradation rates in manure composting or in soils. Half-life estimates ranged from less than four days (in soil, (Sassman and Lee 2007)), to 17 days (manure composting, (Dolliver et al. 2008)), to greater than 70 days (anaerobic manure, (Donoho 1984)). Clearly, the presence of oxygen may determine the speed of degradation of this aromatic molecule.

2.4 Homoacetogenesis

2.4.1 Metabolism

Via the Wood-Ljundahl (acetyl-CoA) pathway (**Figure 2.2**), homoacetogens can use hydrogen and carbon dioxide to produce acetate, though they are capable of using a variety of different one-carbon compounds to form acetate. Of the bacteria that carry out this pathway, the genera *Clostridium* and *Acetobacterium* are the best characterized, but this pathway has been found in many other genera such as *Moorella* and *Sporumosa* (Schiel-Bengelsdorf and Dürre 2012). For the purposes of this literature review, understanding the pathway of homoacetogenesis is important because at the hydrogen partial pressures relevant for efficient *n*-caproate production, homoacetogenesis may also occur and contribute to the control of hydrogen partial pressures in the reactor system. In addition, homoacetogenesis forms acetate, which may be used as a substrate for the reverse beta-oxidation processes.

The Wood-Ljundahl pathway consists of two branches, the carbonyl and methyl branch. In the carbonyl branch, one mole of carbon dioxide is reduced to carbon monoxide. In the methyl branch, another mole of carbon dioxide is reduced in a series of reactions involving six electrons to a methyl group, consuming ATP. The carbonyl and methyl group are then condensed with one mole CoA to form acetyl-CoA and then acetate, which generates ATP (Diekert and Wohlfarth 1994, Ragsdale 2008, Tracy et al. 2012). There is no net ATP production *via* substrate level phosphorylation. However, it has been recently shown that acetogens (*i.e.*, *Acetobacterium woodii* and *Clostridium ljungdahlii*) can link the endergonic reaction catalyzed by the methylene THF reductase enzyme in the methyl branch of the pathway to FBEB and the RNF complex to conserve energy (Hattori 2008, Latif et al. 2014, Poehlein et al. 2012, Tremblay et al. 2013).

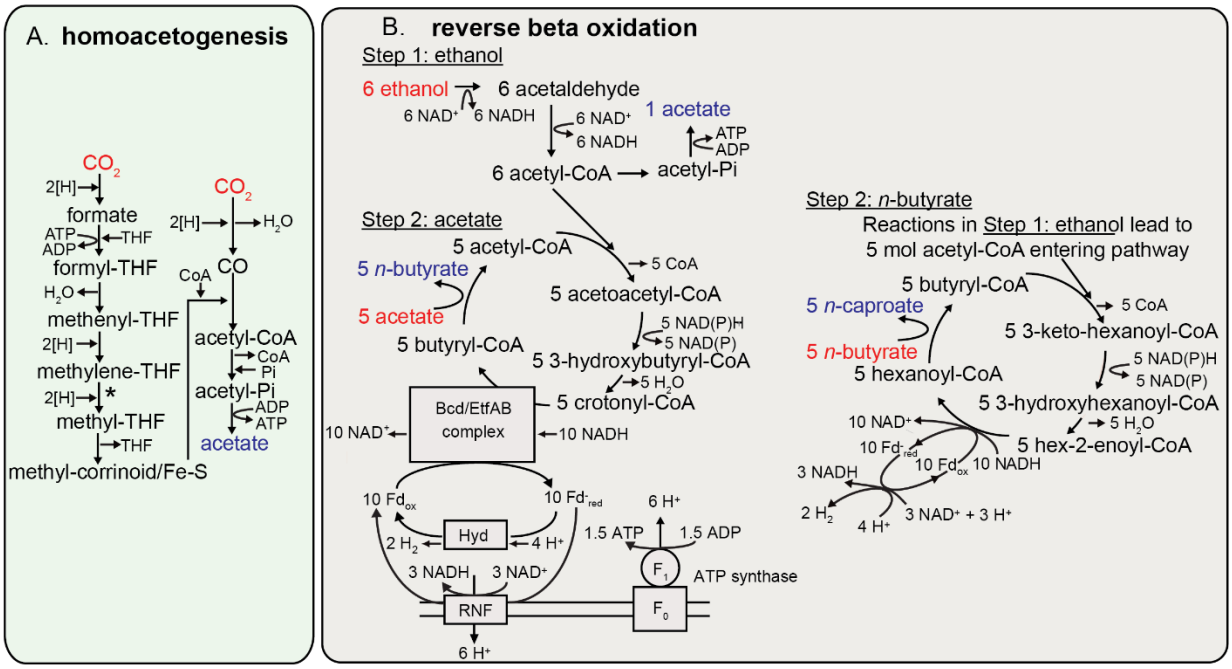

Figure 2.2. Pathways of homoacetogenesis (**A**) and reverse beta-oxidation (**B**) as discussed in this review. For homoacetogenesis, a * has been placed in the pathway where some acetogens have been shown to conserve energy *via* the RNF complex coupled with FBEB. For reverse beta-oxidation, the RNF, ATP synthase, hydrogenase (Hyd), and Bcd/EtfAB complex are shown in detail for the ethanol and acetate to *n*-butyrate pathway and are shown in a simplified format for the ethanol and *n*-butyrate to *n*-caproate pathway (Figure adapted from Seedorf et al. (2008)).

2.4.2 Thermodynamics

The Wood-Ljundahl pathway is reversible, and theoretical work has indicated that the energy coupling sites are the same in both the oxidative and reductive directions (Gonzalez-Cabaleiro et al. 2013). The reductive pathway is more thermodynamically feasible at higher hydrogen partial pressures (Diekert and Wohlfarth 1994), since acetate is formed from hydrogen and carbon dioxide. *Via* thermodynamic calculations, Kleerebezem and Loosdrecht (2010) predicted that at hydrogen partial pressures less than 10^{-2} kPa (*i.e.*, methanogenic conditions) and temperatures greater than 45°C, the reductive pathway would cease to be feasible. At low hydrogen partial pressures, hydrogenotrophic methanogens can outcompete homoacetogens for hydrogen and

carbon dioxide and *vice versa* at higher hydrogen partial pressures (Spirito et al. 2014).

2.4.3 Research and Industrial Applications

The study of the homoacetogenesis is important not only because it plays an important role in carbon cycling in many anaerobic environments (Leang et al. 2013), but also for the potential role of the pathway in the industrial production of acetate, which is an important bulk chemical. Acetate is an important precursor of many industrial chemicals, including polyvinyl acetate. Currently, around 10 million tons of acetate is produced globally per year (Schiel-Bengelsdorf and Dürre 2012), mainly *via* petrochemical methods. The Wood-Ljundahl pathway could be used to fix the carbon dioxide that is commonly lost during industrial fermentation of sugars (Tracy et al. 2012), potentially reducing greenhouse gas emissions and offering a relatively cheap substrate source. The electron donor hydrogen could be derived from the electrolysis of surplus, renewable electric power. However, lower rates of acetate formation *via* homoacetogenesis have been observed in mixed cultures as compared to pure cultures (Spirito et al. 2014).

2.5 Reverse Beta-Oxidation

2.5.1 Metabolism

In reverse beta-oxidation, which is a biological chain-elongation process, a two carbon acetyl-CoA is derived from ethanol and added to a carboxylate, elongating the carbon chain length of the carboxylate two carbons at a time (*i.e.*, acetate to *n*-butyrate, *n*-butyrate to *n*-caproate, and *n*-caproate to *n*-caprylate) (Agler et al. 2011) (**Fig. 2.2, Table 2.1**). Specifically, ethanol is oxidized to acetate, which leads to the formation of NADH and ATP *via* SLP. Acetate is elongated to *n*-butyrate in the cyclic reverse beta-oxidation pathway, which uses NADH for reducing equivalents.

n-Butyrate can be further elongated *via* a similar pathway to *n*-caproate. Hydrogen is also produced in this process. Odd number carbon chains can be elongated *via* this pathway as well (*i.e.*, propionate to *n*-valerate, *n*-valerate to *n*-heptanoate) (Ding et al. 2010, Grootscholten et al. 2013c).

C. kluyveri is one microbe capable of carrying out this metabolic pathway (Agler et al. 2011, Seedorf et al. 2008, Steinbusch et al. 2011, Weimer and Stevenson 2012). A limited number of other microbes have been shown to form *n*-caproate using various substrates, including *Eubacterium alactolyticus*, *Eubacterium biforme*, *Eubacterium limosum*, *Eubacterium pyruvativorans* (which has been shown to take amino acids and form *n*-caproate; (Wallace et al. 2004, Wallace et al. 2003)), *Peptococcus niger,* and *Megasphaera elsdenni* (Ding et al. 2010). For its growth, *C. kluyveri* derives about 30% of its cellular material from carbon dioxide and the rest comes from acetate (Gottschalk 1986). In 2008, the genome of *C. kluyveri* was sequenced revealing that *C. kluyveri* can fix nitrogen and has a highly active sulfur metabolism (Seedorf et al. 2008).

As mentioned previously, it was recently discovered that some anaerobic bacteria, including *C. kluyveri*, utilize FBEB as an alternative energy conservation mechanism in addition to SLP. Seedorf *et al.* (2008) suggest *C. kluyveri* as a model for other anaerobes that can carry out similar processes (Wallace et al. 2004). In *C. kluyveri*, a cytoplasmic enzyme complex (butyryl-CoA dehydrogenase: EtfAB (Bcd/EtfAB)) carries out FBEB by coupling the exergonic reduction of crotonyl-CoA ($E^{\circ\prime}=-10$ mV) to butyryl-CoA by the electron donor NADH ($E^{\circ\prime}=-320$ mV) with the endergonic reduction by NADH of ferredoxin ($E^{\circ\prime}= -410$ mV) (**Fig. 2.1**). The reduced ferredoxin can then be reoxidized to dissipate reducing equivalents as molecular hydrogen in a reaction mediated by a Fe-Fe hydrogenase ($2\ Fd_{red}^{2-} + 4\ H^{+} \rightarrow 2\ Fd_{ox} + 2\ H_2$), which increases the drive for SLP in the oxidative part of the fermentation (*i.e.*, ethanol oxidation). The reduced

ferredoxin can also be reoxidized *via* NAD^+ in the RNF complex ($\Delta E°' = 220$ mV), generating a proton ion gradient that drives the synthesis of ATP *via* ATP synthase (Herrmann et al. 2008). Another enzyme also exists in *C. kluyveri* (the NfnAB complex), which can carry out the following FBEB reaction:

$$Fd_{red}^{2-} + 2\ NADP^+ + NADH + H^+ \rightarrow Fd_{ox} + 2\ NADPH + NAD^+$$

and has been found encoded in other *Clostridium* spp. as well (Wang et al. 2010).

2.5.2 Thermodynamics

Recent calculations have indicated that this pathway is thermodynamically feasible under a range of hydrogen partial pressures (at pH 7, 25°C), which confirms previous experimental work (Gonzalez-Cabaleiro et al. 2013). In this ethanol-acetate fermentation, the ratio of products to substrates is low (**Table 2.1**). Therefore, when the concentrations of all the compounds decrease, the reaction becomes less exergonic. As the available Gibbs free energy decreases, there is a decrease in the amount of ATP that can be generated. This leads to a shift in the enzymes involved in the reverse beta-oxidation pathway and an increasing amount of Fd_{red} is reoxidized *via* the Nfn complex *vs.* the Rnf complex (Buckel and Thauer 2013, Wang et al. 2010). Looking at ethanol and acetate conversion to *n*-butyrate, Buckel and Thauer (2013) determined that when the concentration of products and substrates decreases to one mM, only one ATP can be formed *via* SLP. Therefore, ATP is no longer formed *via* the Rnf complex and ATP synthase, and the Nfn complex is solely responsible for carrying out FBEB. In this situation, Buckel and Thauer (2013) predict that the reverse beta-oxidation pathway shifts from the conversion of six moles of ethanol and four moles of acetate to five moles of *n*-butyrate toward a different stoichiometry: five moles of ethanol and three moles of acetate producing four moles of *n*-butyrate (*i.e.*, from equation a to

b in **Table 2.1**).

Hydrogen can also be used as an electron donor in reverse beta-oxidation (Steinbusch et al. 2011). However, Gonzalez-Cabaleiro et al. (2013) calculated that direct formation of *n*-butyrate from acetate with hydrogen as the electron donor in the reverse beta-oxidation pathway is thermodynamically unfeasible even under high hydrogen partial pressures. The researchers identified a thermodynamic bottleneck in the pathway due to a highly endergonic two acetyl-CoA condensation reaction, which requires unfeasibly low acetyl-CoA concentrations to proceed forward (under evaluated H_2 partial pressures of $1x10^{-3}$, $1x10^{-1}$, 10, 50, and 100 kPa (25°C and pH 7)). *n*-Caproate could be indirectly generated from hydrogen and carbon dioxide *via* homoacetogenesis, acetate reduction to ethanol, and the coupling of ethanol oxidation and reverse beta-oxidation. However, acetate reduction to ethanol would be slow based on kinetics. Ultimately, ethanol is a much better source of reducing equivalents and will lead to increased carboxylate synthesis and biomass growth (Ding et al. 2010).

2.5.3 Pure Culture n-Caproate Production Studies

Early work on this pathway mainly focused on pure cultures of *C. kluyveri*. Barker and Taha (1942) first reported the isolation of *C. kluyveri* from fresh water and marine black mud. In a follow-up study, Barker *et al.* (1945) confirmed *via* labelled carbon experiments that *C. kluyveri* could oxidize ethanol to acetate and couple this to *n*-butyrate and *n*-caproate formation. With a labelled carbon experiment, Tomlinson and Barker (1954) showed that *C. kluyveri* was dependent on carbon dioxide for growth. Much of the early energy metabolism work focused on understanding how *C. kluyveri* conserved energy and generated hydrogen when growing on ethanol and acetate (Jungermann et al. 1969, Smith et al. 1985, Thauer et al. 1968, Thauer et al. 1969). From the early

work on this pathway, it was hypothesized that *C. kluyveri* formed hydrogen from NADH and NAD^+ in a ferredoxin dependent reaction *via* NADH:ferredoxin oxidoreductase. However, it was not understood how this endergonic reaction could proceed (Li et al. 2008). Then, Li *et al.* (2008) discovered the butyryl-CoA dehydrogenase EtfAB complex in *C. kluyveri*, which was capable (*via* FBEB) of coupling this endergonic reduction of ferredoxin with the exergonic reduction of crotonyl-CoA to butyryl-CoA, as was previously discussed. Early researchers also did not understand how, if hydrogen was formed *via* NADH oxidation and ferredoxin reduction, the reduced ferredoxin was reoxidized. In addition, the free energy for *n*-caproate formation ($\Delta G°'$=-180 $kJ \cdot mol^{-1}$) predicted an ATP yield of 2.5 ATP. While it was known that 1 ATP was formed *via* SLP, it was unclear by what mechanism *C. kluyveri* conserved the remainder of its energy (Buckel and Thauer 2013). Sequencing of the genome of *C. kluyveri* demonstrated the existence of a RNF complex (Seedorf et al. 2008). *Via* the RNF complex, *C. kluyveri* can reoxidize the reduced ferredoxin *via* NAD^+, which drives the generation of ATP *via* ATP synthase (Wang et al. 2010). Therefore, it became clear that the ATP yield per reaction was a sum of the ATP synthesis *via* SLP and *via* the RNF complex and ATP synthase *(Herrmann et al. 2008)* (Ding et al. 2010). Additional discoveries, such as that ferredoxin can also be reoxidized *via* protons to form molecular hydrogen with the Fe-Fe hydrogenase and the existence of the Nfn complex in *C. kluyveri* as an alternative to the Rnf complex, further added to our understanding of the metabolism of *C. kluyveri* (Wang et al. 2010).

Some pure culture researchers also examined whether *C. kluyveri* can utilize alternative substrates. Kenealy and Waselefsky (1985) demonstrated that *C. kluyveri* could produce *n*-caproate from propanol and acetate and from succinate and ethanol. More recent work (Jeon et al. 2010, Jeon et al. 2013) has looked at the formation of *n*-caproate from galactitol. Another focus of

pure culture work in recent years has been on genetic engineering. Recently, Dellomonaco et al. (2011) reported success with the engineering of the reverse beta-oxidation pathway into *Escherichia coli*. Finally, as noted before, *n*-caproate production is not exclusive to *C. kluyveri*. For example, some recent pure culture work has focused on the production of *n*-caproate by *E. pyruvativorans*, which is a non-saccharolytic rumen anaerobe (Wallace et al. 2004, Wallace et al. 2003).

2.5.4 Mixed Culture n-Caproate Production Studies

In their review paper on biorefineries, Levy *et al.* (1981) proposed the idea of inhibiting methane production in anaerobic fermentations to promote the formation of medium-chain carboxylates, such as *n*-caproate, and then to extract these carboxylates out of the solution *via* solvent extraction. Smith and McCarty (1989a, b) were the first research group to report low levels of *n*-caproate formation from anaerobic reactor microbiomes (**Table 2.3**). Then, in the 1990's, Holtzapple *et al.* (1999) developed the Mixed Alco process, which focuses on fermentation to form a mixture of carboxylates, which are then concentrated, thermodynamically decomposed to ketones, and subsequently catalytically hydrogenated to a mixture of alcohols. A decade later, Ding *et al.* (2010) examined *n*-caproate formation during mixed culture fermentative hydrogen production *via* experiments and thermodynamic calculations. The first demonstration of high rates of selective *n*-caproate production came when Steinbusch *et al.* (2011) achieved the highest reported *n*-caproate concentration at the time (70 mM; with a production rate of 25 mM C *n*-caproate per day) from synthetic ethanol and acetate. These researchers also reported for the first time the production of *n*-caprylate (an eight-carbon carboxylate) from mixed cultures and hypothesized that it was also formed *via* reverse beta-oxidation. As their system was operated at neutral pH, these researchers

added 2-bromoethanosulfonic acid (BES) to their system, which is an expensive methanogen inhibitor.

Subsequently, Agler *et al.* (2012) used real corn fermentation beer from the corn-to-ethanol industry as substrate (*i.e.*, ethanol and leftover yeast, sugars, and corn grain biomass) and reported achieving *n*-caproate production rates of 108.3 mM C per day and an *n*-caproate specificity of 79% (*i.e.*, carbon in *n*-caproate compared to all fermentation products). The authors avoided BES addition by lowering the pH to 5.5 to inhibit acetoclastic methanogenesis and installing an in-line, membrane-based, liquid-liquid extraction system to prevent buildup of toxic products, based on the earlier work of Wu and Yang (2003). Hydrogenotrophic methanogenesis was not inhibited at pH 5.5 in the reactors. However, low levels of carbon dioxide in the reactor limited the rates of methane formation from hydrogen and carbon dioxide. Methane formation was seen as a beneficial use of the carbon in carbon dioxide, as long as it did not subtract too greatly from the overall carbon flux toward *n*-caproate formation (Agler et al. 2014, Agler et al. 2012a).

Further studies by others have also been carried out without BES addition using various reactor configurations and substrates and many have achieved medium-chain carboxylate production rates comparable to anaerobic digester systems treating solid waste (**Table 2.3**). These studies have looked at using high rate upflow anaerobic filters (Grootscholten et al. 2013b, d, Vasudevan et al. 2014), the production of odd chain carboxylates such as *n*-heptanoate (Grootscholten et al. 2013c), the possibility of two-stage acidification and chain-elongation reactors (Grootscholten et al. 2013a, Grootscholten et al. 2014), and the use of syngas fermentation effluent as a substrate (Vasudevan et al. 2014). Recently, researchers also looked at *n*-caproate production from hydrogen and carbon dioxide and, as expected based on the previous discussion, found a very low volumetric production rate of 1.6 mM $C \cdot d^{-1}$ (Zhang et al. 2013a).

There are still improvements to be made in *n*-caproate production technology before it can be scaled up to full-scale such as improved extraction efficiency, a lower electron donor to *n*-caproate ratio, and a better understanding of the reactor microbiome and what constitutes stability in these systems. The latter two improvements are focused on in this review.

Table 2.3. Review table of studies related to *n*-caproate production with anaerobic reactor microbiomes. In some studies, multiple temperatures, pH, or HRTs were evaluated. The values reported in the table represents the values at which the maximum *n*-caproate concentration and/or production rates were achieved. Definitions of abbreviations used in table: CSTR - continuously stirred tank reactor; UFBR - upflow fixed-bed reactor; ASBR - anaerobic sequencing batch reactor; UAF - upflow anaerobic filter; OFMSW - organic fraction of municipal solid waste.

Substrate	Reactor	Temp (°C)	pH	HRT	Prod. Rate (mM C per day)	Specificity	Max Conc. (mM)	Significance	Source
propionate, ethanol	CSTR	35	7	10 d	NA	NA	5	ethanol perturbation of ADs leads to reduced product formation	(Smith and McCarty 1989a, b)
glucose	UFBR	30	5-6	2-3 h	NA	NA	23	*n*-caproate production during fermentative hydrogen production	(Ding et al. 2010)
acetate, ethanol, H_2	batch	30	7	NA	25.6	62%	70	early study on *n*-caproate, *n*-caprylate formation	(Steinbusch et al. 2011)
corn beer	ASBR	30	5.5	12 d	108.3	79%	54	use of in-line extraction and lower pH (avoid BES addition)	(Agler et al. 2012a)
acetate and ethanol	UAF	30	6.5-7	17 h	812	94%	96	*n*-caproate and *n*-caprylate	(Grootscholten et al. 2013b)

Substrate	Reactor	Temp (°C)	pH	HRT	Prod. Rate (mM C per day)	Specificity	Max Conc. (mM)	Significance	Source
								reported	
carbohydrate wastes	batch	30	4.4	NA	NA	11%	3.3	modified headspace experiments: highest *n*-caproate concentration under 2 bar H_2 headspace	(Arslan et al. 2012, Arslan et al. 2013)
OFMSW, ethanol	batch	30	5-6	NA	36	NA	24	suggests need for two-stage system	(Grootscholten et al. 2013a)
acetate and ethanol	UAF	30	6.5-7.2	Range (4-30 h)	2,886 (4 h HRT)	85% (4 h HRT)	103 (16h HRT)	highest *n*-caproate volumetric production rate achieved	(Grootscholten et al. 2013d)
H_2 and CO_2	biofilm	35	6	NA	1.6	NA	8.5	low yields from H_2 and CO_2	(Zhang et al. 2013a)
OFMSW, ethanol	batch, CSTR	30	6.5-7	11 h	NA	NA	109	Two-stage system	(Grootscholten et al. 2014)
syngas fermentation effluent	AF	30	5.4	14 h	88.5	NA	8.7	syngas fermentation effluent as substrate	(Vasudevan et al. 2014)

2.5.5 Effect of Hydrogen Partial Pressures

Hydrogen partial pressures can affect reactor microbiomes fermenting substrates to *n*-caproate in the following ways: i. by affecting the distribution of the primary fermentation products (*i.e.*, the acetate:*n*-butyrate ratio), which can then affect the amount of ethanol (or other electron donor) needed for chain elongation; ii. hydrogen partial pressures must be high enough to prevent the oxidation of carboxylates and promote formation of more reduced products (**Fig. 2.3**). In the following paragraphs, the role of hydrogen in *n*-caproate formation will be discussed in more detail. In addition, the manipulation of carbon dioxide partial pressures as a way to control hydrogen partial pressures will also be discussed.

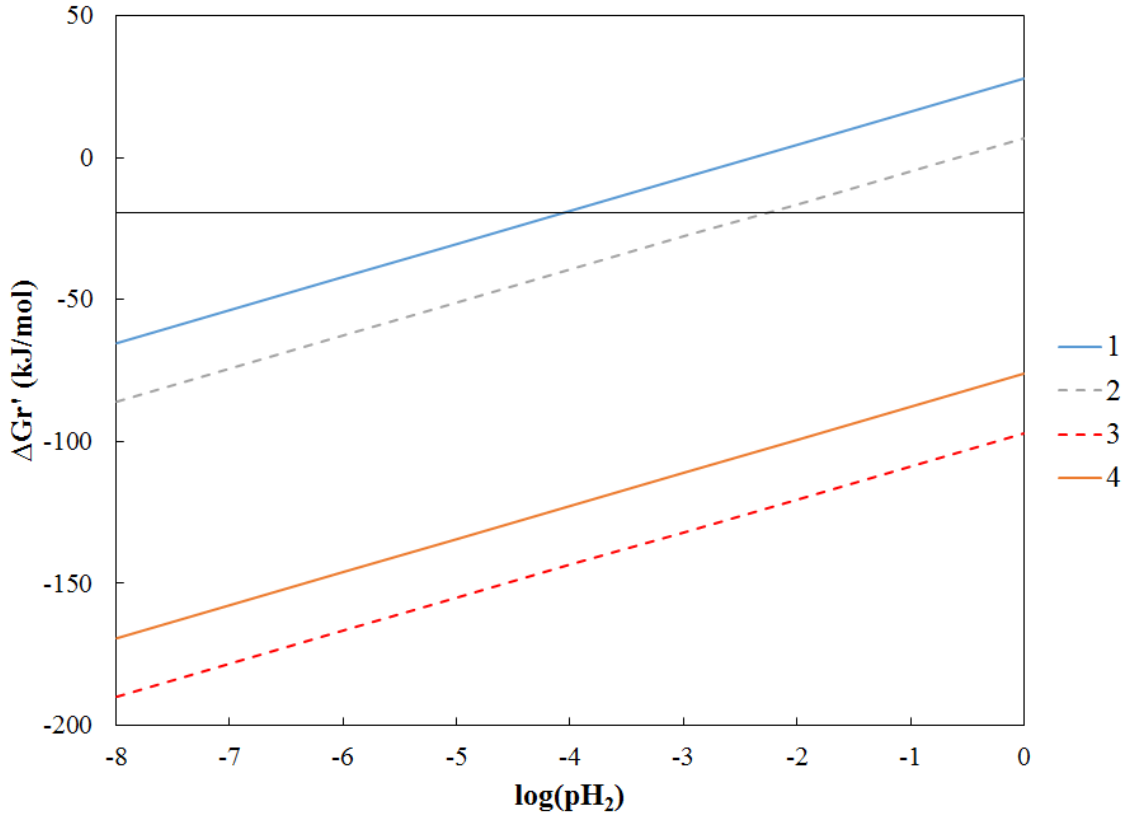

Figure 2.3. Free energy of reaction *vs.* log of hydrogen partial pressure (atm) at 30°C and pH 5.5 for 1. *n*-butyrate oxidation: *n*-butyrate$^-$ + 2 H_2O → 2 acetate$^-$ + H^+ + 2 H_2; 2. Ethanol oxidation:

ethanol + H_20 → acetate$^-$ + H^+ + $2H_2$; 3. *n*-Butyrate formation: 6 ethanol + 4 acetate$^-$ → 5 *n*-butyrate$^-$ + H^+ + 2 H_2 + 4 H_2O; 4. *n*-Butyrate formation *via* the alternative equation published by Buckel and Thauer (2013): 5 ethanol + 3 acetate$^-$ → 4 *n*-butyrate$^-$ + 1 H^+ + 3 H_2O + 2 H_2. *n*-Caproate oxidation to *n*-butyrate and acetate and *n*-caproate formation from ethanol and *n*-butyrate are not shown because their free energy values overlapped almost completely with the lines for *n*-butyrate oxidation (1) and *n*-butyrate formation from acetate and ethanol (3). The concentration of all other substrates and products (besides H^+ and H_2O) were assumed to be 50 mM. The solid black line represents minimum free energy requirement for the reaction of approximately -20 kJ·mol^{-1}.

To understand the effect of hydrogen partial pressures on *n*-caproate production, it is important to first gain an understanding of the production and consumption of hydrogen in these systems. As discussed previously, in anaerobic fermentation systems, the microbial oxidation of alcohols and carboxylates leads to the formation of hydrogen. Specifically, the oxidation of these substrates leads to the formation of electrons, which can be used to reduce protons ($E°'_{H2}$=−414 mV) to hydrogen *via* ferredoxin (E°′~−400 mV, depending on the source) or NADH (E°′=−320 mV), depending on the hydrogen partial pressures in the system (Angenent et al. 2004). Recently, it has been seen that many anaerobes can carry out FBEB to drive formation of hydrogen *via* NADH and Fd_{red}, which would otherwise be endergonic at higher hydrogen partial pressures (Zhang et al. 2013b). Besides the photosynthetic bacteria, *Clostridia* represent the majority of hydrogen producers (Sträuber et al. 2012). In these fermentation systems, the produced hydrogen can be consumed by hydrogenotrophic methanogens, sulfate reducing bacteria, or iron reducing bacteria (among others), depending on the prevailing conditions.

Many previous studies have looked at the effect of hydrogen partial pressures on the distribution of primary fermentation products (*i.e.*, acetate:*n*-butyrate ratio). Though experimental results differ, generally researchers find a shift toward more intermediate product formation (*i.e.*, *n*-butyrate *vs.* acetate) at higher hydrogen partial pressures. From the standpoint of *n*-caproate production and the electron donor (*i.e.*, ethanol) to product ratio, it is desirable to direct primary

fermentation to a lower acetate:*n*-butyrate ratio. *n*-Caproate formation from acetate and ethanol requires more ethanol than *n*-caproate formation *via n*-butyrate and ethanol. A brief overview of these primary fermentation studies will be given here. Previously, Andel *et al*. (1985) found that an increased hydrogen partial pressure led to a decreased acetate:*n*-butyrate ratio in glucose fermentation by *Clostridium butyricum*. Bastidas-Oyanedel *et al*. (2012) also saw a link between changes in headspace gas composition and the products of acidogenesis. Finally, based on a glucose fermentation model including electron bifurcation, Zhang *et al*. (2013b) estimated that the acetate:*n*-butyrate ratio would be less than one if the hydrogen partial pressure was less than ~0.25 atm at pH 5.5. Conversely, Inanc *et al*. (1999) ran an experiment in glucose fed batch CSTRs and concluded that the distribution of volatile fatty acids produced during acidogenesis was mainly affected by population dynamics rather than the partial pressure of hydrogen in the system.

For *n*-caproate production, the partial pressure of hydrogen in the fermentation system must be high enough to prevent oxidation of the carboxylates and direct the fermentation toward more reduced end products (**Fig. 2.3**). Under standard conditions, *n*-butyrate oxidation to acetate and hydrogen is endergonic. According to Harper and Pohland (1986), for physiologically relevant systems, *n*-butyrate oxidation is thermodynamically feasible at hydrogen partial pressures of $\sim 10^{-3}$ atm or below (depending on the exact pH and temperature of the system), while propionate oxidation is feasible at hydrogen partial pressures below $\sim 10^{-4}$ atm. Therefore, hydrogen partial pressures greater than $\sim 10^{-4}$ atm lead to the accumulation of propionic and *n*-butyric acids and inhibit their oxidation by obligate hydrogen producing anaerobes, which is desirable from an *n*-caproate production standpoint. Higher hydrogen partial pressures make it more energetically favorable for microbes to use volatile fatty acids as electron acceptors (Steinbusch et al. 2011).

For *C. kluyveri* metabolism, ethanol oxidation to acetate is coupled to chain elongation of

acetate to *n*-butyrate (or *n*-butyrate to *n*-caproate). Therefore, hydrogen partial pressures can reach higher levels than if ethanol oxidation to acetate was uncoupled. In fact, based on thermodynamics calculations of the author, this pathway should only be unfeasible at hydrogen partial pressures >>1 atm. From previous work on *n*-caproate production (Agler 2011), a hypothesis was developed that, in addition to *C. kluyveri*, there is a consortia of microbes that can carry out chain elongation, involving multiple microbes that can oxidize ethanol to acetate and *Syntrophomonadaceae* spp., which can utilize crotonate or potentially another intermediate metabolite to produce *n*-caproate (McInerney and Wofford 1992), though the hypothesis was not verified. In calculations performed in this literature review (**Fig. 2.3**), it was calculated that at substrate and product concentrations of 50 mM, pH 5.5, and temperature of 30°C, that ethanol oxidation to acetate would be thermodynamically feasible ($\Delta G_r \leq -20$ kJ·mol^{-1}) under hydrogen partial pressures of ~10^{-2} atm. Even if all the *n*-caproate is formed *via* the coupled mechanism used by *C. kluyveri*, some ethanol oxidation to acetate, which does not enter the chain-elongation pathway, can be allowed. As the ethanol is oxidized to acetate and hydrogen, some of the hydrogen will be consumed by hydrogenotrophic methanogens. However, as mentioned before, the carbon dioxide levels are limited in these systems, so eventually the hydrogen levels would build up, inhibiting further ethanol oxidation and shifting the ethanol toward the reverse beta-oxidation pathway and *n*-caproate formation (Agler et al. 2014).

Another gas that has the potential to effect *n*-caproate formation in mixed culture fermentations is carbon dioxide. The addition or removal of carbon dioxide from an anaerobic bioreactor can be used to control hydrogen partial pressures. As discussed previously, hydrogenotrophic methanogens can use hydrogen and carbon dioxide to produce methane. Therefore, supplying carbon dioxide to a reactor should theoretically lower hydrogen partial

pressures, both by decreasing the volume fraction of hydrogen in the headspace and by directing the hydrogen toward hydrogenotrophic methanogenesis. In anaerobic fermentation systems, carbon dioxide is either a substrate in carboxylation reactions or a metabolic product in decarboxylation reactions (Jones and Greenfield 1982). The distribution of carbon dioxide between the headspace and liquid in the reactor is governed by Henry's law and the bicarbonate: carbon dioxide equilibrium.

Carbon dioxide has been shown to have an inhibitory effect on acetogens, which can in turn have an effect on the distribution of fermentation products (Arslan et al. 2012). However, the levels at which carbon dioxide has been found to be inhibitory is much higher than the levels present in efficient *n*-caproate producing bioreactors. For example, Hansson and Molin (1981) observed that 1 bar carbon dioxide partial pressure limited the ability of acetogens to degrade *n*-butyrate to acetate. Similarly, Kim *et al*. (2006) found that carbon dioxide partial pressures above 0.7 bar decreased the biomass activity of acetogens. In the *n*-caproate producing reactor operated in the Angenent Lab, the carbon dioxide partial pressure always remained at levels much lower than what would be considered inhibitory (Agler et al. 2014, Agler et al. 2012a).

Conflicting and few experimental reports exist on the effect of hydrogen on *n*-caproate formation from ethanol and acetate. For example, Arslan *et al*. (2012) imposed a two bar modified headspace (CO_2, H_2, or 50/50 mix of both gases) on reactors treating carbohydrate, protein, or lipid wastes. Contrary to their expectations, they did not see a shift to more reduced compound formation (*i.e.*, *n*-caproate formation) with increased hydrogen partial pressures. In fact, the researchers found that higher hydrogen partial pressures did not limit acetate and *n*-butyrate production. Although two bar hydrogen partial pressure increased total carboxylate production *vs.* the control reactor, it did not result in selective production of a single carboxylate, though it did

result in the highest *n*-caproate concentration observed. The authors hypothesized that either: i. hydrogen was used as an electron donor in their system; ii. mixed cultures are able to tolerate higher hydrogen partial pressures if the hydrogen is not produced *via* fermentation reactions; or iii. the experiments may have needed to be run for a longer time to see a shift toward reduced compounds formation.

2.5.6 Post-Processing and Commercial Use

Commercially, *n*-caproate is produced *via* various chemical methods such as *via* the catalytic oxidation of *n*-hexane (Brink et al. 2000) or from palm oil extraction. Production of *n*-caproate *via* anaerobic microbiomes with organic waste substrates offers a more sustainable alternative to traditional methods. Agler *et al*. (2011) provided a discussion on different post-processing steps to upgrade carboxylates to carbonyls, esters, alcohols, and alkanes. For example, ketonization uses two molecules of carboxylic acids to produce a ketone (e.g., *n*-caproate to undecane, a C11 alkane), which can be further blended into kerosene or diesel (Gaertner et al. 2009, Renz 2005). In addition, *n*-caproate can be used as an animal feed additive, green antimicrobial (Skřivanová and Marounek 2007), or corrosion inhibitor (Grootscholten et al. 2013d).

The need to continuously extract the produced carboxylic acids from the fermentation broth has been previously discussed. Three potential separation methods will be focused on here: liquid-liquid extraction (LLE), ion exchange, and electrodialysis. In LLE, a solvent is used to extract the desired compound from the feed liquid into the solvent (Huang and Ramaswamy 2013). The compound can then be back-extracted into an aqueous solution *via* a pH gradient (Agler et al. 2014). This technology will favor the extraction of more hydrophobic compounds such as *n*-caproic acid (*i.e.*, *n*-caproic acid has a solubility of 10.19 $g \cdot L^{-1}$ in water at 25°C; (CRC 2013)),

from the fermentation broth (Agler et al. 2014, Agler et al. 2012a). Ion exchange resins rely on a stoichiometric process in which certain compounds, such as carboxylic acids, are adsorbed to the resin and removed from the fermentation broth (Huang and Ramaswamy 2013). One drawback is the use of regeneration salts, which creates a large volume of wastewater (Teella et al. 2011). Electrodialysis is a process in which ion exchange membranes are used to selectively extract ions from one solution to another, based on an electrical potential difference. This technology has been applied for several processes such as: i. the separation of acetic acid from acetaldehyde wastewater (Zhang et al. 2011); and ii. the recovery of succinic acid (Huang and Ramaswamy 2013).

2.6 Microbial Ecology Tools and Applications

2.6.1 Overview of Metagenomic Tools

High-throughput sequencing technologies have become increasingly popular in recent years for studies of microbial communities in environments ranging from oceans (Chow et al. 2013) to anaerobic digesters (Nelson et al. 2011, Sundberg et al. 2013). These technologies, along with advances in data analysis tools, permit researchers to carry out high resolution studies examining the changes in microbial communities across spatial or temporal scales. For example, the Illumina MiSeq platform, which is capable of carrying out paired-end 16S rRNA gene sequencing, can generate millions of sequences from a single sample at a relatively low cost (Caporaso et al. 2012, Vanwonterghem et al. 2014b). In 16S rRNA gene sequencing, universal primers targeting the 16S rRNA gene in bacteria and archaea are used to generate amplicons, which are then sequenced *via* platforms such as Illumina MiSeq. As the 16S rRNA gene is found in all bacteria and archaea, sequencing of regions of this gene enables researchers to examine the phylogenetic and taxonomic composition of their samples.

Sequencing of the 16S rRNA gene enables researchers to compare the community composition in samples across temporal or spatial scales, but it does not provide detailed insight into the function and activity of the microbes sequenced in the samples. Often, function can be inferred based on correlations between changes in the microbial community structure and overall function in the study environment (Werner et al. 2011), but to get a more complete picture it is often necessary to turn to other meta-omics techniques such as shotgun metagenomics, metaproteomics, metatranscriptomics, and metabolomics. Two of these meta-omics approaches will be discussed here. Shotgun metagenomics for which the DNA extracted from a sample is sheared into short segments, randomly sequenced (*e.g.*, *via* Illumina HiSeq), and then reassembled into whole or partial genomes, can provide insight into the function of microbes in a community sample. Shotgun metagenomics allows researchers to see what functional genes are present in a sample, but it does not provide insight into whether these genes are being actively expressed. For that, metaproteomics is a useful tool. Metaproteomics involves extracting proteins from a microbial community sample; fractionation, separation, and detection of the proteins (*i.e., via* LC-MS); and, finally, combining the protein data with metagenomic data to characterize the proteins detected (Vanwonterghem et al. 2014b). In one anaerobic digestion study combining metagenomics and metaproteomics, Hanreich et al. (2013) demonstrated that even though metagenomics indicated that methanogens had a relatively low presence in the community, metaproteomics revealed that the methanogens were highly active in the community. Thus, a combination of approaches may be necessary to fully understand the dynamics of the microbial community in question.

As mentioned above, these high throughput techniques generate massive amounts of sequencing data, which are then combined with metadata from the study environment (*i.e.*,

operating and environmental parameters in the case of an anaerobic digester system). To analyze this data, it is essential to have a firm grasp of the different microbial ecology and multivariate statistical analyses and tools available, which I will briefly cover here. Traditional microbial ecology analyses for 16S rRNA gene sequencing data involves looking at the alpha (within sample) and beta (between samples) diversity of the microbial community samples from a certain environment. Metrics that exist to measure the alpha diversity of microbial community samples typically account for the richness (*i.e.,* number of species) and/or evenness (*i.e.,* relative abundance of species) in a sample. For example, the Gini coefficient is one metric that can be used to assess the evenness of a microbial community, with a Gini coefficient of 0 representing a perfectly even community and a Gini coefficient of 1 representing an uneven community dominated by one species (Wittebolle et al. 2009). The Shannon diversity coefficient is another alpha diversity metric that takes into account the richness and evenness of the community in question (Shannon 1948). In terms of beta diversity, one metric that can be used is the Unifrac distance, which is a measure of the phylogenetic distance between two sample communities (Lozupone et al. 2010). To visually display beta diversity data, researchers can carry out unconstrained ordination (*e.g.*, principal coordinates analysis (PCoA)). Through constrained ordination (*e.g.,* distance-based redundancy analysis (db-RDA)), metadata can be used to attempt to find a link between changes in microbial community composition and function. QIIME (Quantitative Insights Into Microbial Ecology) is one open-source software pipeline designed to analyze high-throughput sequencing data (Caporaso et al. 2010). Packages in R and other software packages can also be used to probe further into the data *via* machine learning (*e.g.*, PAMR for nearest shrunken centroid analysis (Hastie et al. 2013) or Orange for linear discriminant analysis (Demšar et al. 2013)) and network analyses (*e.g.*, Cytoscape (Cline et al. 2007, Shannon et al. 2003)). For example, network analysis

can be used to examine co-occurrence patterns in microbial communities (Barberán et al. 2011).

Sequencing data from shotgun metagenomics studies can be compared to the MG-RAST database (Meyer et al. 2008) to identify clusters of orthologous groups (COGs) of proteins. In addition, *de-novo* assembly (*e.g.*, *via* Velvet (Zerbino and Birney 2008)) can be performed to group the shorter sequences in the sample into larger contigs that can then be functionally annotated based on reference databases. A review by Thomas et al. (2012) and a more recent review by Sangwan et al. (2016) provide details on the steps to take in collecting, processing, and analyzing shotgun metagenomics data. Protein data from the LC-MS metaproteomics analysis can also be combined with the *de-novo* assembly data to identify the protein sequences and gain a picture of their relative abundance (Hanreich et al. 2013).

In sequencing studies, an important focus is often on linking operating and environmental parameters with microbial community changes. In Werner et al. (2011), multivariate statistical tools were used to uncover a stronger relationship between community structure and function *vs.* environment in their anaerobic digesters. They suggested that these tools open up the possibility to engineer anaerobic digester communities with improved performance. In their *n*-butyrate reactor study, Agler et al. (2012b) used similar tools to see how substrate pretreatment operating conditions and history affected performance parameters. They were able to statistically separate changes in microbial composition from performance changes. In their review paper, Vanwonterghem et al. (2014b) cited the importance of the use of these tools to link community composition and dynamics with operating conditions and performance parameters to gain insight into the metabolic functions of different microbes involved in anaerobic digestion. These tools have also been applied in natural systems. In a study by Wang et al. (2013), the authors examined deterministic *vs.* stochastic processes in various natural environments, such as lake water and soils.

The authors found that unmeasured environmental variables rather than stochastic processes were driving some of the observed changes in the studied microbiomes. Hartman et al. (2014) also used these tools to examine the correlations between changes in the microbiomes and physical parameters for compacted soils.

2.6.2 Stability and Disturbances

High throughput sequencing studies allow researchers to examine whether broad temporal trends exist for microbial communities and whether changes in microbial communities are predictable. A greater understanding of how a normally functioning community changes with time will enable better identification of when a community is undergoing a disturbance and better predictions of how a community will change in response to a disturbance (Shade et al. 2013). Therefore, it is necessary to both examine microbial communities under typical conditions and under perturbations. Researchers are increasingly attempting to examine the response of these communities to disturbances (*e.g.*, the resistance, resilience, and functional redundancy (Allison and Martiny 2008, Shade et al. 2012)), to examine what constitutes a stable community, and to quantify both the temporal variability and phylogenetic turnover in a stable environment. Bioreactors are controlled engineered systems which can be easily manipulated. Therefore, they are ideal systems in which to test microbial ecology theories related to stability and the effect of perturbations (Hai et al. 2014).

Researchers can quantify the response of a community to a disturbance *via* terms such as resistance, resilience, and functional redundancy (Allison and Martiny 2008, Shade et al. 2012). Resistance is the degree to which a community remains unchanged in response to a disturbance; resilience is the degree to which the community returns to its initial state following a disturbance;

and functional redundancy refers to the ability of one or more microbial populations to fill the ecological niche of other, negatively-impacted microbial populations in a reactor system. In their brewery wastewater anaerobic digester study, Werner et al. (2011) found that methanogenic communities were resilient and quickly returned to their original composition and functions following disturbances. In bioreactor studies by Fernandez et al. (1999, 2000) and Hashsham et al. (2000), the researchers found that the dominant microbes under steady-state conditions were often outcompeted during perturbed conditions by previously low abundant microbes. This indicates that the insurance hypothesis, which states that species with ecological functions similar to another species may increase in abundance while another declines during a disturbance, plays an important role in bioreactor systems (Botton et al. 2006). In some situations, microbial communities never return to their initial state following a disturbance and instead shift to an alternative stable state (Shade et al. 2011, Stein et al. 2013).

Several studies have examined what contributes to a functionally stable and robust community in bioreactor systems. Early important work in this area was carried out by Fernandez et al. (1999, 2000) and Hashsham et al. (2000) on methanogenic bioreactors subjected to glucose shocks. These researchers found that the bioreactors with microbial communities that were less structurally stable and diverse were more functionally stable, with more functionally redundant organisms. In contrast, in a study of mesophilic anaerobic digesters treating brewery wastewater, Werner et al. (2011) linked increased functional stability to increased community structure evenness (as measured by the Gini coefficient), which confirmed previous work by Wittebolle et al. (2009).

Fewer studies have been conducted on the stability of the microbiomes involved in other chain-elongation processes such as reverse beta-oxidation. In their study of *n*-caproate producing

anaerobic bioreactor, Agler et al. (2012a), observed that improved performance was correlated with an increasingly uneven and less diverse (lower richness) community in their system. In addition, in a study by Agler et al. (2012b) of thermophilic *n*-butyrate producing bioreactors, the authors observed a relatively uneven community structure compared to the community of the robust mesophilic anaerobic digester systems of Werner et al. (2011) which was discussed earlier. Though Agler et al. (2012b) found a more uneven community, the researchers did find evidence of functional redundancy in their system, which may be a key component of stability.

Though anaerobic digesters have been found to be functionally stable, many studies have reported that the microbial communities of the digesters can be dynamic and variable during 'steady-state' performance (Vanwonterghem et al. 2014b). Therefore, it is useful to gather an understanding of what constitutes stable, steady-state variability in community composition and what indicates a disturbance. One way of quantifying change in a microbial community with time is to look at the phylogenetic turnover and temporal variability (Wang et al. 2013). In Shade et al. (2013), the authors carried out a meta-analysis of changes in the microbial communities of different environments with time (ranging from the infant gut to brewery wastewater anaerobic digesters). The authors quantified the time decay in these communities and found that the community turnover was generally quite slow. They observed that temporal variability in both alpha and beta diversity was generally similar for samples from the same or similar environments. In their meta-analysis, brewery wastewater and soil samples were generally less variable than other community types, especially *vs.* communities undergoing primary succession such as the infant gut. Another interesting aspect of their meta-analysis is that many of the relationships related to species turnover in microbial systems back up what has previously been discovered in macro ecology studies of plants and animals (Van Der Gast et al. 2008).

2.6.3 Reactor Microbiomes Involved in Carboxylate Platform Processes

As mentioned above, these microbial ecology tools have been applied to carboxylate platform systems to examine the microbial communities present. For example, researchers have applied these tools to anaerobic digesters. Riviere et al. (2009) used Sanger sequencing to investigate the core microbiome of seven different full-scale anaerobic digesters at municipal wastewater treatment plants. Several studies have noted the effect of substrate type on the microbial community of anaerobic digesters (Sundberg et al. 2013, Wirth et al. 2012, Zhang et al. 2014). Few studies have conducted a time-course analysis of the changes in the microbial community of anaerobic digesters, though the work of Town et al. (2014) and Werner et al. (2011) should be noted. These studies, along with previous cultivation-based studies, have given insight into what microbes dominate in different anaerobic digester systems and the role these different organisms play in anaerobic digestion of organic wastes. For example, one recent metagenomic study found a link between *Clostridia* populations and hydrogenotrophic methanogens, suggesting a possible syntrophic association (Vanwonterghem et al. 2014a). Bacteriodetes and Firmicutes phyla are often found to be dominant in anaerobic digester systems and are linked to the breakdown of polysaccharides and fermentation (Town et al. 2014, Vanwonterghem et al. 2014a, Vanwonterghem et al. 2014b, Wirth et al. 2012).

Though few next-generation-based sequencing studies have looked at the effect of antibiotic disturbances on anaerobic digester communities, several studies exist on the effect of antibiotics on the human gut (Antonopoulos et al. 2009, Dethlefsen et al. 2008, Stein et al. 2013). Stein et al. (2013) carried out a time-series study using mice on the effect of antibiotics on the human gut. Generally, the authors note that broad spectrum antibiotics act by removing highly

abundant species, which can allow less abundant, antibiotic resistant bacteria to dominate. The authors found that antibiotic administration led to a shift in the gut community to an alternative stable state and the community did not return to its initial state.

Recently, studies have sought to characterize the microbiome involved in medium-chain carboxylate production, mainly *via* 16S rRNA gene sequencing (Andersen et al. 2017, Hollister et al. 2010, Hollister et al. 2012, Steinbusch et al. 2011). For example, in the study by Agler et al. (2012), the authors carried out 16S rRNA gene sequencing and found five abundant OTUs, of various taxonomies, whose relative abundances that were significantly positively correlated ($r>0.8$ and $p<0.05$) with increased *n*-caproate production, including *C. kluyveri,* as well as a member of the *Ruminococcaceae* family. In more recent studies by Kucek et al., the authors employed 16S rRNA gene sequencing to characterize the microbiomes in a bioreactor converting *l*-lactate to *n*-caproate (Kucek et al. 2016b); in a bioreactor converting wine lees to *n*-caproate and *n*-caprylate (Kucek et al. 2016a); and in a bioreactor converting ethanol and acetate to primarily *n*-caprylate (Kucek et al. 2016c). In addition, some recent work has been performed in isolating chain-elongating microbes. For example, Weimer et al. (2012) isolated a *C. kluyveri* strain capable of medium-chain carboxylate production from the bovine rumen, Jeon et al. (2016) isolated a *Megasphaera elsdenii* strain capable of medium-chain carboxylate production from bovine rumen, and a *Clostridium cluster* IV member in the family Ruminococcaceae was isolated from lactate-fed bioreactor by Zhu et al. (2017).

2.7 Conclusions

In this review, I have covered the metabolism, thermodynamics, and practical applications of several carboxylate platform pathways: anaerobic digestion, homoacetogenesis, and reverse beta-

oxidation. The effect of perturbations on two of these pathways (*i.e.*, the effect of an antibiotic on anaerobic digestion and the effect of hydrogen partial pressures on reverse beta-oxidation) was discussed. Finally, I reviewed the high throughput sequencing and microbial ecology analyses tools that will enable researchers to carry out in-depth, time-series studies of these systems. For a thorough understanding of the effect of perturbations on these carboxylate platform systems, it is important to tie together an understanding of the metabolism, thermodynamics, system performance, and microbial community dynamics for these pathways.

CHAPTER 3

RESISTANCE AND ADAPTATION TO THE ANTIBIOTIC MONENSIN BY THE ANAEROBIC DIGESTION MICROBIOME

Adapted from: Spirito, Daly, Werner, and Angenent. *Submitted to Frontiers in Microbiology*, May 2017

3.1 Abstract

The antibiotic monensin is fed to dairy cows to increase milk production efficiency. A fraction of this monensin is excreted in the cow manure. Previous studies have found that cow manure containing monensin can negatively impact the performance of anaerobic digesters, especially upon first introduction. Few studies have examined whether the anaerobic digester microbiome can adapt to monensin during the operating time. Here, we conducted a long-term, time-series study of four lab-scale anaerobic digesters fed with cow manure. We examined changes in both the microbiome composition and function of the anaerobic digesters when subjected to the dairy antibiotic monensin. Two anaerobic digesters received manure from monensin-dosed cows; while the other two anaerobic digesters received manure from control cows with monensin added directly to the digester. In our digesters, the monensin was not rapidly degraded under anaerobic conditions. The anaerobic digesters subjected to manure from the monensin-dosed cows exhibited relatively small changes in microbiome composition and function, because of the relatively low monensin concentrations of less than 0.2 $mg \cdot L^{-1}$ in the manure. In the anaerobic digesters where monensin was directly added, higher monensin concentrations of up to 5 $mg \cdot L^{-1}$ drove major changes in the anaerobic digester microbiome composition and function. Rapid introduction of

monensin to one of these anaerobic digesters led to a significant decrease in methane yield. Conversely, more gradual addition of the same concentrations of monensin to the other anaerobic digester led to the apparent adaptation of the anaerobic digester microbiome to the high monensin concentration. A member of the candidate OP11 (Microgenomates) phylum arose in the anaerobic digester and appeared to be functionally redundant with certain Bacteroidetes phylum members, which had previously dominated the anaerobic digester. In all digesters, some common shifts were observed in the primary-fermenter and syntrophic populations with monensin addition. This study demonstrated how anaerobic digester microbiomes resist and adapt to the monensin antibiotic.

3.2 Introduction

Time-series studies of environmental microbiomes allow microbial ecologists to both characterize the functional stability of the microbiome and to explore how that microbiome shifts in response to a disturbance. To examine the concept of stability in environmental microbiomes subjected to a disturbance, ecologists have recently sought to quantify the response of a population subjected to a given disturbance in terms of resistance (*i.e.*, the ability of a population to remain unchanged), resilience (*i.e.,* the return or recovery of a population to its pre-disturbance abundance and function), and redundancy (*i.e.*, the replacement of negatively impacted populations by others that fill its ecological niche) (Allison and Martiny 2008). These environmental studies have ranged from lake microbiomes subjected to water column mixing (Shade et al. 2011) to the microbiomes of forest soils subjected to compaction (Hartmann et al. 2014). An increased understanding of the temporal variability inherent in microbial communities is vital to predict how communities will respond to disturbances (Shade et al. 2013). However, studies looking at stability as related to natural environments often are faced with numerous variables that are difficult to control. In

contrast, controlled engineered environments, such as anaerobic digesters, represent ideal systems to carry out time-series studies examining the effect of disturbances on the microbiome (de los Reyes Iii et al. 2015, Vanwonterghem et al. 2014a, Werner et al. 2011).

Anaerobic digestion is a biotechnology platform that is used to treat organic wastes and produce methane. It is a mature technology, which is applied around the world on farms and at municipal and industrial wastewater treatment plants. In the controlled environment of an anaerobic digester, a complex microbial food web breaks down organic wastes *via* hydrolysis, acidogenesis, acetogenesis, and methanogenesis to the main end product gaseous methane. The process stability relies on syntrophic interactions between microbes carrying out different components of the anaerobic digestion pathway (Schink 1997). Process instability can result when certain members of the anaerobic food web are inhibited and can no longer carry out their roles. Inhibition may be caused by disturbances, such as substrate overloading or antibiotic addition (Chen et al. 2008). To maintain process stability in the face of disturbances, the anaerobic digestion microbiome relies on a high degree of functional redundancy. Recent studies have noted that, after subjected to changes in environmental or operating conditions, the anaerobic digester microbiome often recovers its function before it returns, if ever, to its original composition (De Vrieze et al. 2017, Langer et al. 2015). In anaerobic digestion, multiple populations are often able to carry out the same function in the digester, which leads to a flexible community structure and increased process stability (Fernandez et al. 2000). Increased understanding of how anaerobic digester microbiomes respond to disturbances is important from both a microbial ecology perspective, as well as a process operating perspective.

Introduction of an inhibiting compound, such as the antibiotic monensin, can result in digester disturbance. Monensin is a monovalent carboxylic ionophore antibiotic that is given to

dairy cows to increase their milk production efficiency or to treat coccidiosis, which is a parasitic infection (Simjee et al. 2012). Monensin alters the fermentation pathways in the cow rumen by inhibiting the precursors of methanogenesis. A previous study by Chen and Wolin (1979) demonstrated that monensin favors the proliferation of propionate-producing bacteria, such as the lactate fermenter *Selenomonas ruminantium*, and inhibits hydrogen-, formate-, and acetate-producing bacteria, such as *Ruminococcus albus*, *Ruminococcus flavefaciens*, and *Butyrivibrio fibrisolvens,* leading to decreased methane production in the cow rumen. In a later study by Dennis et al. (1981), the authors demonstrated that other bacteria that are capable of fermenting lactate to propionate, such as *Megasphaera elsdenii*, *Anaerovibrio lipolytica*, and *Veillonella alcalescens*, as well as succinate producing bacteria, are resistant to monensin. Similar to other ionophore antibiotics, monensin accumulates in the cell membrane of bacteria, which interferes with the ion gradients needed to transport nutrients and generate proton motive force. The affected bacterium diverts its ATP away from other essential cell processes to try to restore the normal ion gradient, and as a result growth ceases (Simjee et al. 2012). In general, Gram-negative bacteria are considered to be more resistant to monensin than Gram-positive bacteria due to their outer lipopolysaccharide layer (Simjee et al. 2012). However, the division is not always clear-cut. A pure culture study by Simjee et al. (2012) on three strains of Gram-positive bacteria isolated from cattle rumen (*i.e.*, *Clostridium perfringens*, *Enterococcus faecium*, and *Enterococcus faecalis*) indicated that these bacteria can develop monensin resistance through altered cell membrane characteristics, such as thickening of the extracellular polysaccharide (glycocalyx) layer or the cell wall. Similarly, Rychlik and Russell (2002) observed adaptation to monensin by a strain of the Gram-positive, amino acid fermenting bacterium *Clostridium aminophilum*. Conversely, some Gram-negative bacteria, such as certain strains of *Prevotella bryanttii* (Callaway and Russell 1999)

and *Prevotella ruminicola* (Morehead and Dawson 1992), are sensitive to monensin and require an adaptation period.

A previous study found that up to 50% of the monensin antibiotic fed to beef cattle was excreted in the manure (Donoho et al. 1978). Previous studies and on-farm experiences have shown a decrease in methane production upon first introduction of monensin or monensin-laden manure to anaerobic digesters (Sun et al. 2014, Thaveesri et al. 1994, Varel and Hashimoto 1981, Zitomer 2007). These studies indicated that monensin indirectly inhibits methanogenesis by depleting the precursor acetic acid. However, a six-month study by Varel and Hashimoto (1982) indicated the capability of an anaerobic digester to adapt to monensin during the operating period. Here, only the concentration of monensin in the cow feed was reported, whereas the concentration of monensin in the adapted digester was not reported. A lack of long-term studies in this area has limited understanding on how to prevent possible digester failure or decreased performance. Recent advances in high-throughput sequencing have enabled researchers to investigate the effect of various antibiotics on anaerobic digester microbiomes (Miller et al. 2016, Mustapha et al. 2016). A recent study used 454 pyrosequencing of the 16S rRNA gene to look at the effect of monensin on microbial populations in the cow rumen (Kim et al. 2014). However, few studies exist examining how the anaerobic digester microbiome resists and adapts to the antibiotic monensin.

Here, a long-term time-series study (383 days) was carried out to investigate the effect of monensin on the microbiome composition and dynamics as well as the performance of four, lab-scale anaerobic digesters carrying out anaerobic digestion of dairy cow manure. These anaerobic digesters were exposed to different disturbance regimes by varying monensin concentrations in the digester feed, either by the addition of manure from monensin-dosed cows or *via* manure from control cows with monensin directly added. Previous studies (Vanwonterghem et al. 2014a,

Werner et al. 2014) have demonstrated that replicate anaerobic digesters follow similar performance and microbiome composition patterns during the operating period. Therefore, in this study, replicate anaerobic digesters were not employed. The overall objective of this present study was to correlate digester performance data with microbiome composition. Specifically, we studied whether: 1) the microbiome became resistant to low concentrations of monensin (less than 0.2 $mg \cdot L^{-1}$) in the manure excreted by monensin-dosed cows; 2) the microbiome became resistant to high concentrations of monensin (of up to 5 $mg \cdot L^{-1}$) directly dosed to the digester, following either a rapid or slow introduction schedule; and 3) the shifts in the anaerobic digester microbiome were associated with adaptation to higher monensin concentrations.

3.3 Materials and Methods

3.3.1 Monensin Dosing to Cows

Seven dairy cows (Cornell University Teaching and Research Center, Dryden, NY) were fed the same diet except that three cows (monensin-dosed cows) received a ration top-dressed with monensin (Rumensin® 90 Premix, Elanco Animal Health, Indianapolis, IN, mixed with corn grain), while four cows (control cows) received no monensin (ration top-dressed with corn grain without Rumensin). The monensin-dosing rate in the top-dress was increased every two weeks for a period of two months, following a typical schedule for introducing monensin to dairy cow diets. The monensin concentration in the top-dress was later measured by liquid chromatography with post-column derivatization, as described by Coleman et al. (1997) (Covance Laboratories, Greenfield, IN) and the monensin-dosing rates were calculated to be: 194, 320, 432, and 546 $mg \cdot d^{-1}$ on a *per* cow basis.

3.3.2 Anaerobic Digester Operation

The manure from the two groups of dairy cows was collected at the end of each monensin dose period to obtain five batches of manure: M0, M200, M300, M400, and M500, corresponding to the 0, 194, 320, 432, and 546 $mg \cdot d^{-1}$ monensin-dosing rates, respectively (M0 was collected from the control cows; while M200-M500 was collected from the monensin-dosed cows). Four, 4.5-L continuously stirred anaerobic digesters were inoculated with a mix of ~3.5-L of digester sludge collected from an anaerobic digester treating municipal wastewater located at the Ithaca Area Wastewater Treatment Facility (Ithaca, NY), ~0.1-L of centrifuged solids from a lab-scale continuously stirred anaerobic digester, and ~0.9-L of blended upflow anaerobic sludge blanket (UASB) anaerobic digester granules from a brewery treatment plant (Budweiser Anheuser-Busch InBev, Baldwinsville, NY). The four anaerobic digesters were then semi-continuously fed (every two days) with manure at a target organic loading rate (OLR) of 2 g volatile solids $(VS) \cdot L^{-1} \cdot d^{-1}$ and a 25-day hydraulic retention time (HRT). We operated the anaerobic digesters at a temperature of 37+1°C.

The four anaerobic digesters are referred to as the Low A, Low B, Fast, and Slow anaerobic digesters, and were all initially fed with the control (M0) manure for a period of 202 days (startup and period 1 [P1]). The anaerobic digesters were then subjected to different monensin-dosing strategies throughout periods 2-4 (P2, P3, and P4) (**Table 3.1**). The Low A and Low B anaerobic digesters were fed manure from monensin-dosed cows from Day 203 (start of P2) throughout P4 (feed-dosed manure; M200-M500). Low A and Low B anaerobic digesters followed similar schedules to each other, except that the final fed-dosed manure batch (M500), which contained the highest concentration of excreted monensin, was introduced earlier into the Low A anaerobic digester than the Low B anaerobic digester (Day 245 *vs.* Day 267, respectively) (**Table 3.1**). The

Fast and Slow anaerobic digesters received manure from the control cows (M0) to which a 0.1-mL volume of a prepared monensin solution (monensin reference standard from Elanco Animal Health, Greenfield, IN, which was dissolved in methanol to achieve the targeted monensin concentration) was directly added (direct-dosed manure). The Fast anaerobic digester was fed direct-dosed manure from Day 203 (P2) until the anaerobic digester was shutdown on the final day of P3 (**Table 3.1**). The Slow anaerobic digester initially acted as control anaerobic digester and was fed control cow manure until Day 306 (P3). After this, it was fed direct-dosed manure until the end of P4 (**Table 3.1**). To account for the possible effect of adding methanol in the substrate, 0.1 mL of methanol (with or without monensin) was added to all of the substrates in the study.

Table 3.1. Monensin concentrations in the influent (Inf.) and the effluent (Eff.) of the anaerobic digesters. Manure batch that was fed is indicated for each time period. Influent values are based on dilution (manure was diluted before feeding). M0+ or M0++ indicate that monensin was added directly to control cow (M0) manure. M200, M300, M400, and M500 refer to the four batches of manure collected from the monensin-dosed cows (corresponding to the 194, 320, 432, and 546 $mg \cdot d^{-1}$ monensin-dosing rates, respectively). P1 to P4 refers to the four main periods in the study. Detection limit was 0.10 $mg \cdot L^{-1}$ for monensin measurements. The Low B anaerobic digester started M500 manure batch on Day 267 and the Low B anaerobic digester was discontinued on Day 355 of the operating period. The Fast anaerobic digester was discontinued on Day 306 of the operating period.

		Low A			**Low B**			**Fast**			**Slow**		
			Monensin Concentration ($mg \cdot L^{-1}$)			**Monensin Concentration ($mg \cdot L^{-1}$)**			**Monensin Concentration ($mg \cdot L^{-1}$)**			**Monensin Concentration ($mg \cdot L^{-1}$)**	
	Days		**Inf.**	**Eff. [day]**		**Inf.**	**Eff. [day]**		**Inf.**	**Eff. [day]**		**Inf.**	**Eff. [day]**
Start-up	1-114	M0	0	na	M0	0	na	M0	0	na	M0	0	na
P1	115-202	M0	0	nd [201]	M0	0	nd [201]	M0	0	nd [201]	M0	0	nd [201]
P2	203-216	M200	0.09	na	M200	0.09	na	M0+	0.09	na	M0	0	na
	217-230	M300	0.12	na	M300	0.12	na	M0+	0.12	na	M0	0	na
	231-244	M400	0.15	nd [231]	M400	0.15	nd [231]	M0+	0.15	nd [231]	M0	0	nd [231]
	245-264	M500	0.19	na	M400	0.15	na	M0+	0.19	na	M0	0	na
P3	265-280	M500	0.19	0.12 [267]	M500	0.19	nd [267]	M0++	1	0.11 [267]	M0	0	nd [267]
	281-306	M500	0.19	0.14 [301]	M500	0.19	0.12 [301]	M0++	5	2.29 [301]	M0	0	nd [301]
		M500	0.19	na	M500	0.19	na	M0++	5	2.73 [306]	M0	0	na

		Low A			Low B			Fast			Slow		
			Monensin Concentration ($mg \cdot L^{-1}$)			Monensin Concentration ($mg \cdot L^{-1}$)			Monensin Concentration ($mg \cdot L^{-1}$)			Monensin Concentration ($mg \cdot L^{-1}$)	
	Days		**Inf.**	**Eff. [day]**		**Inf.**	**Eff. [day]**		**Inf.**	**Eff. [day]**		**Inf.**	**Eff. [day]**
P4	307-320	M500	0.19	na	M500	0.19	na	stopped			M0++	1	na
	321-338	M500	0.19	0.11 [331]	M500	0.19	0.11 [331]	stopped			M0++	2	0.81 [331]
	339-352	M500	0.19	na	M500	0.19	na	stopped			M0++	3	na
	353-366	M500	0.19	0.13 [365]	M500	0.19	0.13 [355]	stopped			M0++	4	1.80 [365]
	367-383	M500	0.19	0.14 [383]	stopped			stopped			M0++	5	2.55 [383]

**na=measurement not available; nd=monensin not detected*

3.3.3 Analytical Methods

Anaerobic digester performance and stability parameters (*i.e.,* pH, total volatile fatty acids (tVFAs) concentrations, individual volatile fatty acids (iVFAs) concentrations, total alkalinity, volatile solids (VS) and total solids (TS) concentrations, soluble COD concentrations, total ammonium concentrations, biogas production rate, and biogas composition) were monitored routinely. In addition, the manure substrate was characterized (*i.e.,* pH, gross energy, tVFAs and iVFAs concentrations, VS and TS concentrations, soluble and total COD concentrations, total alkalinity, and total ammonium concentrations). All analyses were carried out using the Standard Methods for the Examination of Water and Wastewater, 21st ed. (Eaton et al. 2005), unless noted otherwise. Biogas composition was measured using a gas chromatograph (SRI 8610C, SRI Instruments, Torrance, CA), which was equipped with a thermal-conductivity detector (TCD) under isothermal conditions (*i.e.,* 105°C) and a packed column (0.3-m HaySep-D packed Teflon; Restek, Bellefonte, PA) with helium as a carrier gas. iVFA concentrations were measured using a gas chromatograph (HP Hewlett Packard 5890 Series II), which was equipped with a flame-ionization detector (FID) with a ramp temperature program (initial temperature 70°C for 2 min; temperature ramp 12°C per min to 200°C; final temperature 200°C for 2 min), and a capillary column (NUKOL, Fused Silica Capillary Column, 15 m × 0.53 mm × 0.50 μm film thickness; Supelco Inc., Bellefonte, PA). The injection port was set to 200°C and the detector to 275°C. The gross energy of the manure was quantified *via* bomb calorimetry analysis by an independent laboratory (Dairy One, Ithaca, NY), using an IKA C2000 basic Calorimeter System (IKA Works, Inc., Wilmington, NC). Effluent samples (200-mL volume) were collected from the anaerobic digesters monthly beginning during P1 and stored in a -23°C freezer for later analysis of monensin concentration. As with the monensin concentration in the topdress described above, liquid chromatography with post-column

derivatization was used to measure monensin concentrations in the manure substrate and anaerobic digester effluent samples. Loss-on-drying analysis (105°C) was used to measure the moisture content of the manure and effluent samples. The monensin concentration in the wet manure and digester samples, as well as the moisture content of the samples, was then used to calculate the concentration of monensin in the samples on a dry matter basis. Monensin and loss-on-drying analyses were performed by Covance Laboratories (Greenfield, IN).

3.3.4 Biomass Sampling

Biomass samples for 16S rRNA gene sequencing analysis were collected from: 1) the hindgut of the dairy cows *via* rectal grab sampling; 2) the manure substrates used for feeding the anaerobic digesters; and 3) the anaerobic digesters. The hindgut biomass samples (184 samples) were collected from control cows and monensin-dosed cows approximately weekly during a three-month period (*i.e.,* for one month *prior* to the start of monensin dosing of the cows and for the two months during which a subset of the cows were subjected to monensin dosing). Duplicate samples were collected from each of the combined manures that were used as substrates for the anaerobic digesters (10 samples). Biomass samples were also collected from the anaerobic digesters on an approximately weekly basis (154 samples). Samples were centrifuged and 2-3 g of the solids were stored at -80°C for 16S rRNA gene sequencing analysis. In addition, we report details of the specific operating periods in which the samples were collected (**Table A1.1)**.

3.3.5 DNA Extraction, Amplification, and Sequencing

DNA was extracted from the cow hindgut, manure, and anaerobic digester biomass samples using the PowerSoil® DNA Isolation Kit (MoBio, Carlsbad, CA, USA). DNA extracted from the cow

hindgut biomass samples was sent to the Earth Microbiome Project at the University of Colorado Boulder for further sample processing (*i.e.,* PCR amplification *via* universal primers targeting the V4 region of the 16S rRNA gene – 515F forward primer and 806R reverse primer, amplicon cleanup, and Illumina HiSeq 2000 sequencing). Details of the sample processing can be found at www.earthmicrobiome.org (Gilbert et al. 2010).

For the cow manure and anaerobic digester samples, we employed a modified version of the Earth Microbiome Project protocol (Gilbert et al. 2010). The modified protocol has been outlined previously by Regueiro et al. (2015), with the exception that in this study 30 PCR cycles were used instead of 25. As in the study by Regueiro et al. (2015), we performed duplicate PCR reactions of the extracted DNA samples and pooled the resulting amplicons *prior* to sequencing. Samples were sent for paired-end sequencing (2x250bp) on the Illumina MiSeq platform (Illumina, San Diego, CA, USA) at the Cornell University Biotechnology Resource Center (Ithaca, NY, USA).

3.3.6 Sequencing Data Analysis

Paired-end reads were joined and then further processed *via* the Quantitative Insights into Microbial Ecology platform (QIIME 1.7) (Caporaso et al. 2010). Quality filtering was performed using the default values in QIIME with the exception that the minimum acceptable Phred quality score was set to 25. After demultiplexing, closed reference OTU picking with the default uclust method (Edgar 2010) was used to group sequences into operational taxonomic units (OTUs) at 97% identity. Taxonomy was assigned with the RDP classifier (Wang et al. 2007) using the Greengenes database (May 2013) for representative sequences selected for each OTU (McDonald et al. 2012). This resulted in 8330 OTUs (~20% of initial 11.8 million quality filtered sequences

were discarded because they did not match any sequences in the reference database). Singletons were removed from the dataset resulting in 6694 OTUs (mean sequences per sample: 52,000). Alpha diversity was analyzed *via* the observed species (*i.e.*, richness) and Gini coefficient metrics (*i.e.*, unevenness; where 0 is equivalent to perfectly even, and 1 is equivalent to uneven with one abundant OTU and all other OTUs as singletons). Ten rarefactions at a depth of 10,300 sequences per sample were performed and collated. The weighted and unweighted Unifrac distance metrics (Lozupone et al. 2006) were used to analyze beta diversity at an even sampling depth of 10,300 sequences per sample. Principal coordinates analysis (PCoA) was used to visualize the differences in community between the samples.

All amplicon and metadata has been made public through the QIITA data portal (qiita.microbio.me) under study number 1621 for the cow hindgut samples and study number 10560 for the cow manure and anaerobic digester biomass samples. Sequences were also submitted to EBI under the following accession number ERP017357.

3.3.7 Statistical Analysis

Constrained ordination, specifically distance-based redundancy analysis (db-RDA), was carried out using the capscale function in the package vegan in R (Oksanen et al. 2015). For the distance-based redundancy analysis, analysis of variance (ANOVA) and the variance inflation factor (VIF) were used to select the environmental and functional parameters that best recreated the sample clustering observed in the PCoA plots. The variables included in the analysis were: bicarbonate alkalinity, soluble COD concentrations, total ammonium concentrations, monensin concentrations in the substrate, specific biogas production rate, methane yield, pH of the effluent and substrate, VS and TS concentrations, and tVFA concentrations. Spearman's rank correlation coefficient was

used in R (Hmisc package) to examine the correlation between monensin substrate concentrations and the relative abundance of the OTU populations. Only OTUs that reached at least 1% relative abundance in any one anaerobic digester sample were considered in the correlation analysis. Correlations were considered significant for $p<0.001$ ($\rho<-0.5$ for negative correlations and $\rho>0.5$ for positive correlations). Statistical analyses of the environmental data were performed using the Tukey HSD model for comparing multiple means by pairwise comparisons in RStudio (v0.96.316).

3.4 Results

3.4.1 Minor Changes in Manure Composition Following Monensin Dosing to Cows

The monensin concentration in the wet manure from the monensin-dosed cows in this study ranged from 0.239 ± 0.009 $mg \cdot L^{-1}$ at the lowest monensin-dosing rate of 194 $mg \cdot d^{-1}$ to 0.543 ± 0.011 $mg \cdot L^{-1}$ at the highest monensin-dosing rate of 546 $mg \cdot d^{-1}$ (**Table A1.2**). We calculated a monensin excretion rate of 13% (R^2=0.82) based on the linear relationship that was observed between the measured monensin concentration in the consumed feed and in the manure, on a dry matter basis in $mg \cdot kg^{-1}$ (**Fig. A1.1**). The different manures fed to the anaerobic digesters were diluted by varying amounts of tap water to achieve the target organic loading rate, which decreased the monensin concentrations in the manure substrate (**Table 3.1**). We observed some differences in the physical and chemical properties of the manures from the monensin-dosed cows *vs.* the control cows, which persisted after the manure was diluted (**Table A1.2**). Total alkalinity and total ammonium concentrations, as well as the pH, were significantly higher ($p<0.05$) and the tVFAs concentration were significantly lower ($p<0.05$) for the majority of monensin-dosed cow manures *vs.* the control cow manure (**Table A1.2**). Total alkalinity, total ammonium concentration, and pH

generally increased with monensin-dosing rate (**Table A1.2**). However, the monensin-dosed cow manures were collected at later time points than the control cow manure, which was only collected at a single time point. Therefore, natural changes in the cow metabolism during our experimental period may also have contributed to the observed differences in the manures. Bomb calorimetry analyses revealed no significant differences in the gross energy of the manures ($p=0.296$) (**Table A1.2**). No clear patterns emerged in the acetate to propionate ratio in the manures (**Table A1.2**).

We observed some differences in the in-between sample diversity (β diversity) for the microbiomes of the monensin-dosed cow *vs.* control cow samples, although more work would be required to be conclusive. A PCoA of unweighted Unifrac distances between cow hindgut samples exhibited clustering of the monensin-dosed cow samples *vs.* the control cow hindgut samples (the monensin-dosed cow samples are all in left half of **Fig. A1.2A**). However, this clustering could not be exclusively attributed to the monensin-dosing rates, because the cow hindgut samples clustered strongly based on the individual cow they came from. Importantly, the samples still clustered within the individual cow whether they were taken two weeks before or during monensin dosing for each of the monensin-dosed cows (**Fig. A1.2A**). We also sequenced the combined manure substrates, which were fed to the anaerobic digesters (M0, M200, M300, M400, and M500). The dominant taxa families observed in the manure substrates are reported in the SI (**Table A1.3**). The PCoA (unweighted Unifrac) of the manure substrates showed a difference between the combined control manure substrate and the monensin manure substrates microbial communities, though, the small sample size should be noted (**Fig. A1.2B**). To conclusively isolate the different monensin-dosing rates as the driver for microbiome dynamics, we would need to repeat the experiment at a larger scale.

3.4.2 Monensin Affects Performance of Anaerobic Digesters

Monensin was not rapidly degraded in the anaerobic environment of the digesters (**Table 3.1**). At the relatively low concentrations of monensin in the influent, which was 0.19 $mg \cdot L^{-1}$ for Low A and Low B during the periods with the highest feed-dosed manure substrate M500 (Day 245 to 383 and Day 267 to 355, respectively), a considerable proportion (~74% and ~68%, respectively) remained in the effluent at the end of these periods (**Table 3.1**). Similarly, at the relatively high concentration of monensin in the Fast and Slow anaerobic digesters, which we dosed with up to 5 $mg \cdot L^{-1}$ monensin (between Days 281 to 306 and Days 367 to 383, respectively), ~54% and ~51%, respectively, remained in the effluent (**Table 3.1**). This result builds upon a previous, shorter-term study by Varel et al. (2012) in which mesophilic laboratory anaerobic digesters were fed with manure from cattle dosed with monensin. After a period of 28 days, the concentration of monensin in the digesters in that study had only decreased by 8%.

The monensin-dosed manure negatively impacted methane yield in the Low A and Low B anaerobic digesters even though the monensin concentration in the manure substrate never exceeded 0.2 $mg \cdot L^{-1}$ (**Fig. 3.1A**). During P2 and P3, when the Low A and Low B anaerobic digesters received the manure from the monensin-dosed cows, they exhibited a slight decrease in average methane yield compared to the control anaerobic digester. The decrease in methane yields was not significant for P2 ($p=0.581$), while it was for P2 ($p<<0.001$). Specifically, we observed a decrease of 3.2% during P2 and a decrease of 14.7% during P3 for the Low A anaerobic digester. For the Low B anaerobic digester, we observed a decrease of 1.9% during P2 and a decrease of 9.9% during P3 (**Fig. 3.1A, Table A1.6**). During these periods, the tVFAs concentration in the anaerobic digesters remained low, reaching an average of 155±29 mg $Ac \cdot L^{-1}$ in the Low A anaerobic digester and 162±27 mg $Ac \cdot L^{-1}$ in the Low B digester during P4 (**Table A1.4, A1.5**).

Thus, it seems unlikely that that the lower methane yields were due to an inhibition of acetogenesis or methanogenesis by the relatively low monensin concentrations that were present in the manures. If this had occurred, we would have anticipated a corresponding build-up of other short-chain volatile fatty acids, such as propionate and *n*-butyrate, as has been described in other papers on anaerobic digester inhibition (Schink 1997).

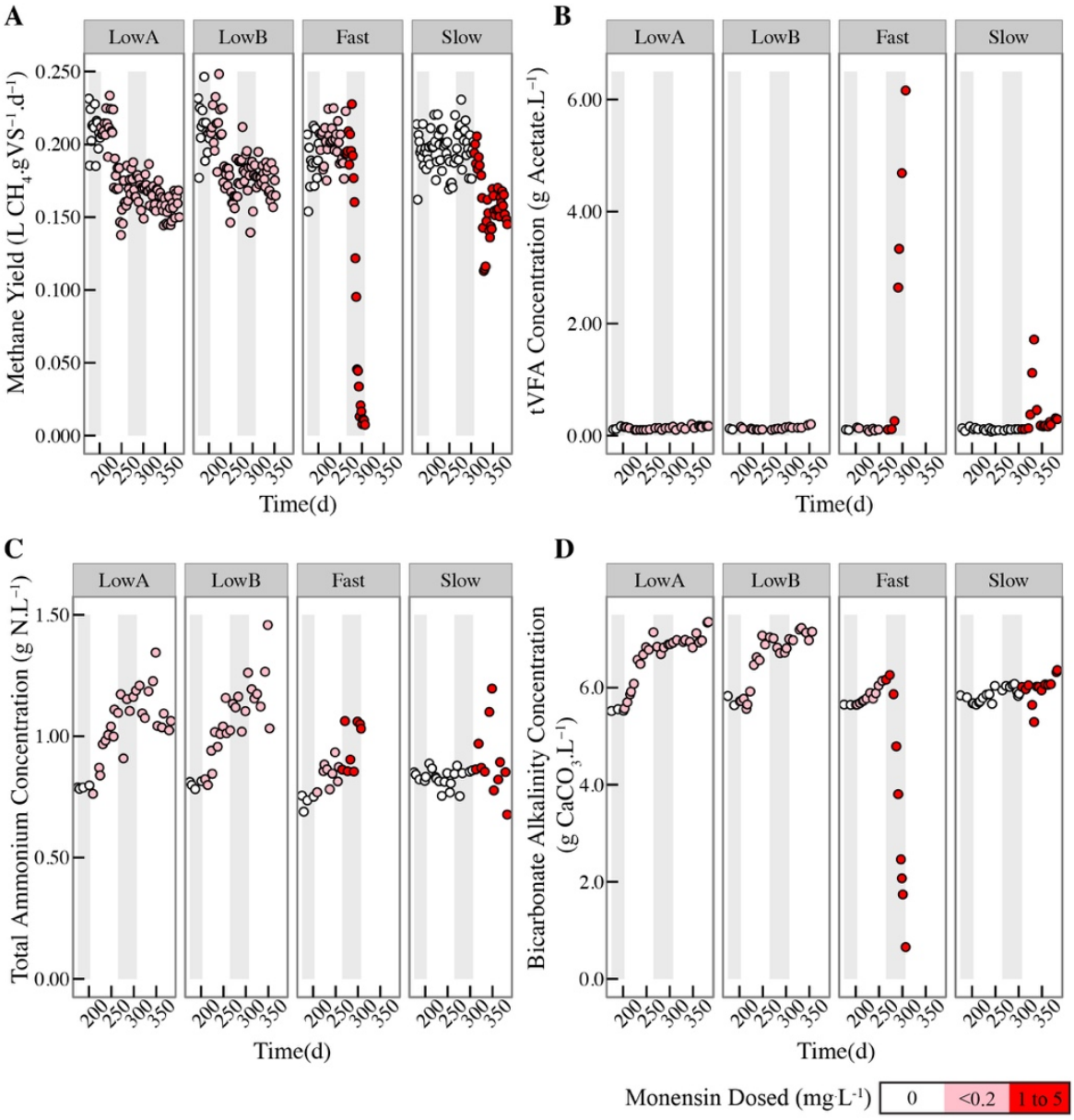

Figure 3.1. Performance parameters for the four anaerobic digesters from Day 175 to the end of the operating period for: **A**) specific methane yield; **B**) tVFAs concentration; **C**) total ammonium concentration; and **D**) bicarbonate alkalinity concentration. Points are colored on a color scale corresponding to the concentration of monensin dosed to the anaerobic digester at that time point.

The decreases in methane yields after the addition of monensin were more pronounced for the Fast and Slow anaerobic digesters compared to the Low A and Low B anaerobic digesters (**Fig.**

3.1A), and coincided with rapidly increasing tVFAs concentrations (**Fig. 3.1B**). For the Fast anaerobic digester, after we increased the monensin dose from 1 to 5 $mg \cdot L^{-1}$ on Day 281 during P3, we observed a rapid decrease in methane yield and an accumulation of tVFAs (>2 g $Ac \cdot L^{-1}$ on Day 291 and reaching a maximum of 6.2 g $Ac \cdot L^{-1}$ on Day 306) (**Fig. 3.1A-B**). The increased VFA concentrations point toward the inhibition of acetogens by high concentrations of monensin. The accumulation of VFAs resulted in a drop of the pH level to 5.81 on Day 306 (data not shown) and a further decrease in the methane yield to only ~0.01 L $CH_4 \cdot g\ VS^{-1} \cdot d^{-1}$ was observed after which the operating period was terminated due to unstable conditions (**Fig. 3.1A**). In the period (P3) *prior* to monensin addition to the Slow anaerobic digester, the average methane yield was 0.20±0.01 L $CH_4 \cdot g\ VS^{-1} \cdot d^{-1}$ (**Table A1.6**). During P4, monensin was dosed to the anaerobic digester starting at 1 $mg \cdot L^{-1}$ and incrementally increased to 5 $mg \cdot L^{-1}$ (**Table 3.1**). During this period, methane yield initially declined but then began to recover (**Figure 3.1A**), resulting in an average methane yield of 0.16±0.02 L $CH_4 \cdot g\ VS^{-1} \cdot d^{-1}$ for P4 (**Table A1.6**). We observed only a temporary increase in the tVFA concentrations in the Slow anaerobic digester when the drop in the methane yield was most severe (**Fig. 3.1A-B**). In both anaerobic digesters, detectable concentrations of other VFA species besides acetate were observed upon initial introduction of high concentrations of monensin (*i.e.,* propionate and *n*-butyrate in the Fast anaerobic digester and propionate in the Slow anaerobic digester; **Table A1.4**). As mentioned earlier, monensin did not rapidly degrade in the anaerobic digesters, resulting in similar concentrations in both anaerobic digesters (**Table 3.1**). Therefore, the main difference between the Fast and Slow anaerobic digesters was the rate at which the concentration of monensin was increased from 1 to 5 $mg \cdot L^{-1}$ (**Table 3.1**). While the Fast anaerobic digester failed, the Slow anaerobic digester partially recovered its performance ~40 days after the first introduction of the monensin antibiotic, as indicated by the methane yield data (**Fig. 3.1A**).

In the Low A and Low B anaerobic digesters, total ammonium concentrations increased as the monensin concentration in the manure substrate from the monensin-dosed cows increased (**Fig. 3.1C and Table A1.5**). However, this could also be attributed to the fact that total ammonium concentrations in the manure from the monensin-dosed cows increased as the monensin dose to the cows increased (**Table A1.2**). Bicarbonate alkalinity concentrations were higher for the Low A and Low B anaerobic digesters compared to the control anaerobic digester after the switch to the monensin-dosed cow manure substrate (**Fig. 3.1D**). This was anticipated because ammonia ions contribute to bicarbonate alkalinity (Graef and Andrews 1974). Even though performance changes were observed between the anaerobic digesters, biomass concentrations (TS and VS), biomass conversion efficiencies, and other stability parameters were similar between all the anaerobic digesters (**Table A1.5 and A1.6**).

3.4.3 Monensin Changes Composition of Anaerobic Digester Microbiomes

For two periods *prior* to the addition of monensin (Startup and P1) (**Table 3.1**), each of the anaerobic digester microbiomes exhibited a similar evolution from the inoculum (Day 0) community (**Table A1.7)** toward a community that was dominated by the phyla Bacteroidetes (61.5±2.7%), the WWE1 candidate division (12.4±1.8%), and Firmicutes (10.8±0.4%) (on Day 201 in **Fig. A1.3A-D**). The richness and unevenness (α diversity) of the anaerobic digester communities changed similarly for all four anaerobic digesters between the startup period and P1 (**Fig. A1.4** and **A1.5**). We observed a considerable decrease in the average observed species (richness) between the startup period and P1 (**Fig. A1.4A-D**), while the Gini coefficient (unevenness) increased between these periods (**Fig. A1.5A-D**). We did not find a statistical difference ($p>0.05$) between the microbiomes of the four anaerobic digesters in terms of the

richness and unevenness measurements during the startup period and P1. The mean weighted Unifrac distance (measurement of β diversity) between each set of reactor samples for these periods was not significantly different ($p>0.05$, based on the Tukey HSD model for comparing multiple means by pairwise comparisons), indicating a similar degree of divergence of the microbiomes during the pre-monensin periods (**Fig. 3.2A**). In addition, during P2-P4 when monensin was present in most of the anaerobic digesters (including the digesters with high monensin concentrations), we did not find significant changes for the richness and unevenness measurements for the anaerobic digester samples with monensin addition (at significance levels of $p<0.05$) (**Fig. A1.4** and **A1.5**).

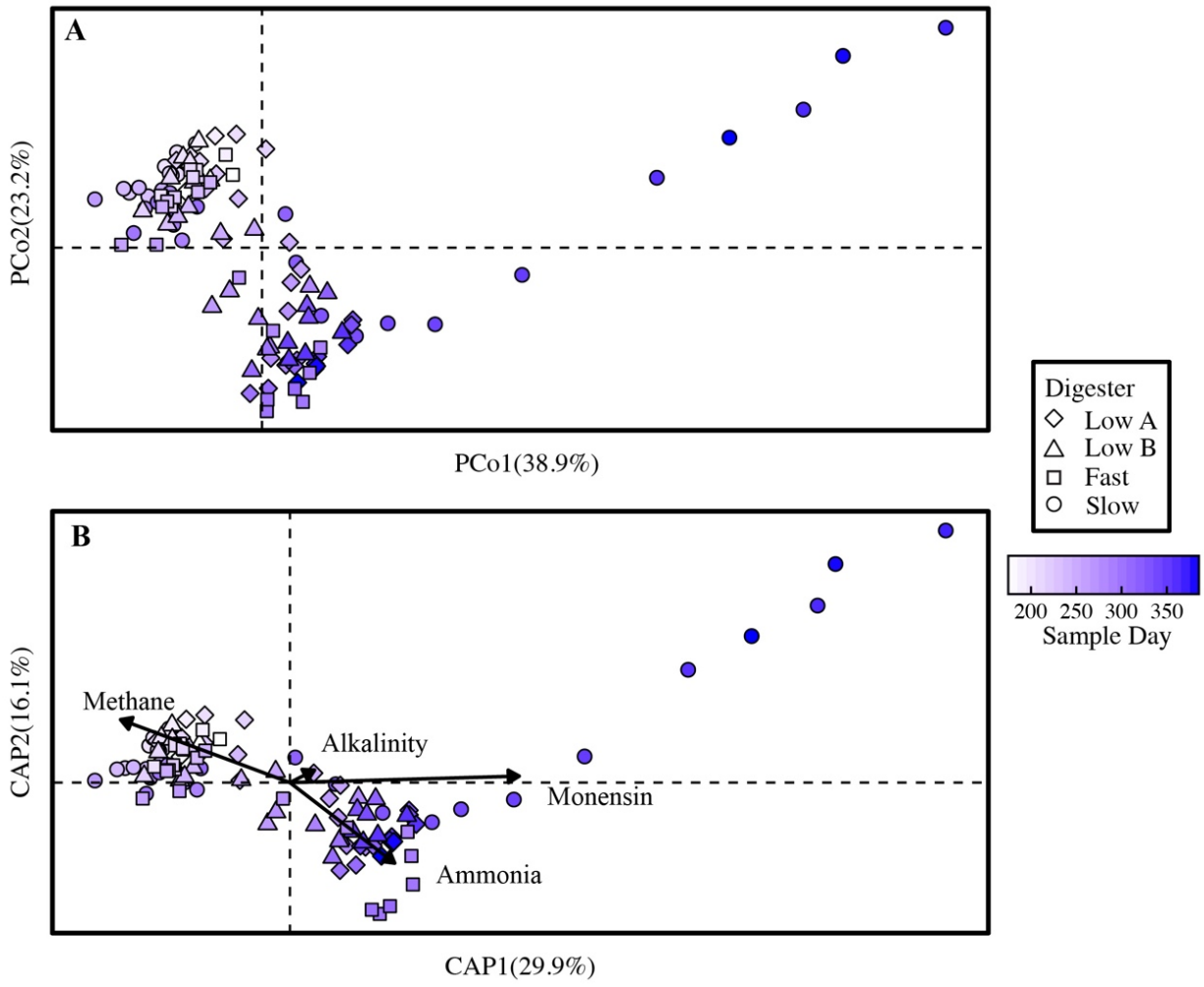

Figure 3.2. β-diversity of all anaerobic digester microbiome samples from Day 175 to the end of the operating period: **A**) principal coordinates analysis (PCoA) plot based on weighted Unifrac distance; and **B**) distance-based redundancy analysis (db-RDA) plot showing the four measured parameters (Methane=methane yield; Alkalinity=bicarbonate alkalinity concentration, Ammonia=total ammonium concentration, Monensin=monensin concentration in the substrate) that best explained the variation observed in the PCoA plot. In both the PCoA and db-RDA plots, the points are colored on a gradient scale representing time.

We did, however, observe changes in the β diversity of the anaerobic digester microbiomes with monensin addition. The Slow anaerobic digester showed the greatest divergence from its initial anaerobic digester community as indicated by the divergence of the Slow digester microbiome away from the other digester samples along the PCo1 axis of the PCoA, which was

based on the weighted Unifrac distance metric (**Fig. 3.2A**). All anaerobic digesters subjected to monensin saw a larger divergence from their initial anaerobic digester communities than the control anaerobic digester (**Fig. A1.6A**). Furthermore, according to the results of a distance-based redundancy analysis (db-RDA or capscale analysis in R), four different environmental and functional parameters best explained the variation seen in the anaerobic digester communities: 1) total ammonium concentrations in the anaerobic digester (environmental); 2) the monensin concentration entering in the substrate (environmental); 3) bicarbonate alkalinity (environmental); and 4) the measured methane yield (function) (**Fig. 3.2B**).

3.4.4 Specific Taxa Shifts in Anaerobic Digesters that Were Exposed to Monensin

Some common patterns were observed in terms of the taxonomic composition of the anaerobic digesters. The phylum of Bacteroidetes (mostly the order of Bacteroidales) was dominant for the majority of the study in all four anaerobic digesters. Dominant OTUs within the order Bacteroidales underwent a notable succession event corresponding with the start of monensin addition in all anaerobic digesters (**Fig. 3.3**; **Table A1.8** and **A1.9**). Besides Bacteroidales OTUs, we found that the relative abundances of other bacterial populations were either positively or negatively correlated with the monensin concentration dosed in all the anaerobic digesters. OTUs belonging to the genus *Pelotomaculum* and the genus *HA73* within the family Dethiosulfovibrionaceae were positively correlated, while other OTUs belonging to the family of Ruminococcaceae, the genus *Syntrophus*, and the genus *Candidatus Cloacomonas* were negatively correlated to the monensin concentration dosed (**Table A1.8** and **A1.9**). Two OTUs, which belong to the genus *T78* within the family Anaerolinaceae, were found to be positively correlated with monensin concentration dosed in only the Fast and Slow digesters (high monensin dose) (**Table**

A1.8).

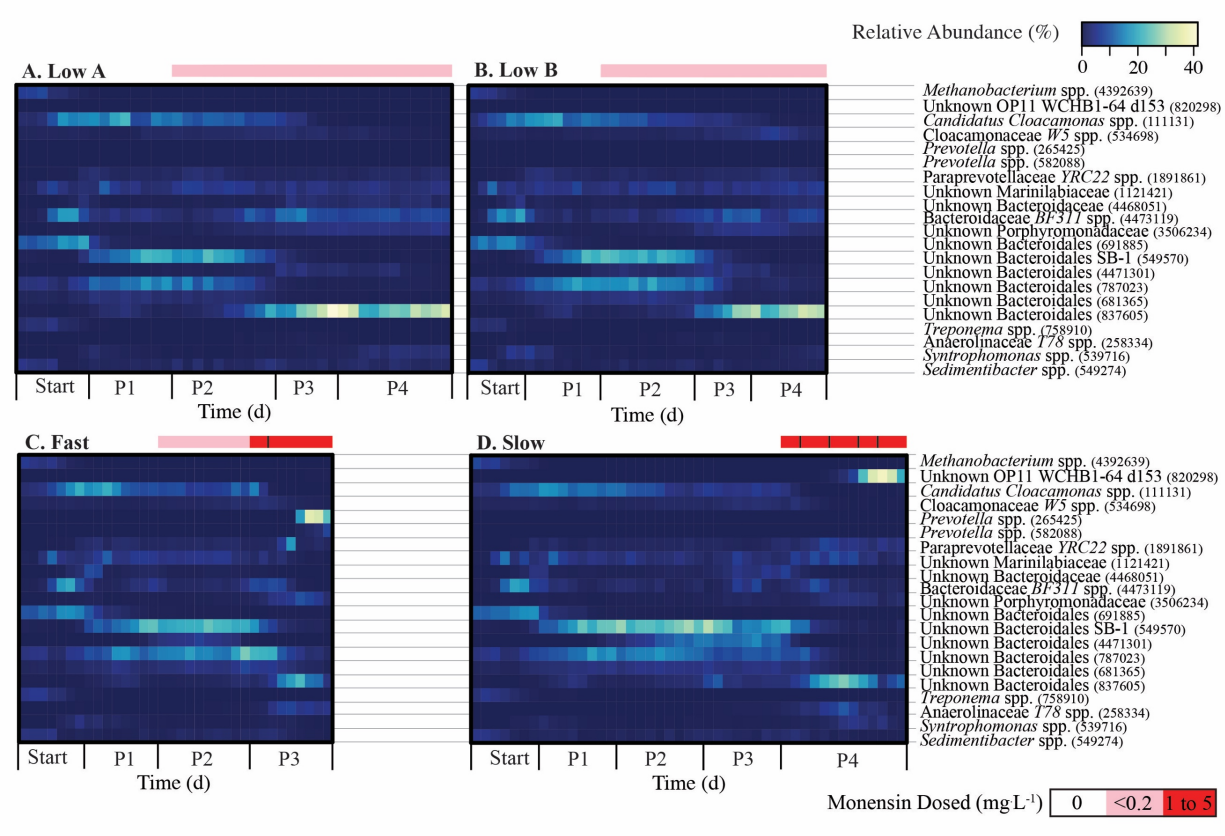

Figure 3.3. Heatmaps representing relative abundance of major OTUs in anaerobic digester samples during the entire sampling period for: **A**) Low A anaerobic digester; **B**) Low B anaerobic digester; **C**) Fast anaerobic digester; and **D**) Slow anaerobic digester. OTUs represented here reached at least 5% relative abundance in any one anaerobic digester sample. Lowest level taxonomic data, as well as the OTU ID number, is provided. The sidebar color scale represents the monensin-dosing rate to the anaerobic digester. For the Fast digester (C), the black line on the red sidebar indicates when the monensin dose was directly increased from 1 to 5 $mg \cdot L^{-1}$. For the Slow digester (D), the black lines on the red sidebar indicate when the monensin dose was increased (from 1 to 2 $mg \cdot L^{-1}$, from 2 to 3 $mg \cdot L^{-1}$, from 3 to 4 $mg \cdot L^{-1}$, and from 4 to 5 $mg \cdot L^{-1}$).

Two different, previously rare OTUs rapidly rose to high levels of relative abundance for the Fast and Slow anaerobic digester during the disturbed conditions. For the Fast anaerobic digester, an OTU (ID 265425) belonging to the genus of *Prevotella* (phylum of Bacteroidetes -

order of Bacteroidales) rapidly increased in abundance, from not detectable on Day 257 (P2 – *prior* to high concentrations of monensin) to a maximum of 35.1% on Day 299 (P3 – high monensin) (**Fig. 3.3C**). This OTU was found at very low abundance levels in the biomass samples from the other anaerobic digesters (**Fig. 3.3A-B,D**). For the Slow anaerobic digester, an unknown OTU (ID 820298) belonging to the order d153 in the phylum of OP11 (Microgenomates) rapidly rose in abundance from <0.01% on Day 301 (pre-monensin) to a maximum of 38.3% on Day 369 (P4 – high monensin) (**Fig. 3.3D**). The increase in this OP11 OTU, which was one of only two OP11 OTUs found in the anaerobic digester and represented the majority of the phylum of OP11, coincided with a decrease in the overall relative abundance of the phylum of Bacteroidetes in the Slow anaerobic digester (**Fig. A1.3D**). The shift in the Bacteroidetes and OP11 phyla OTU populations drove the divergence of the Slow anaerobic digester microbiome, which can be seen along the PCo1 axis of the principle coordinate analysis plot (**Fig. 3.2A**). The OP11 OTU also increased in the Fast anaerobic digester after the dose with the high concentrations of monensin, but to a much lower extent than in the Slow anaerobic digester. The OTU increased from not detectable on Day 257 (P2 – *prior* to high concentrations of monensin) to a maximum of 0.3% on Day 306 (P3 – high monensin) (**Fig. 3.3C**).

The dominant methanogenic genera in the seven Day 0 (inoculum) samples that were sequenced were the acetoclastic methanogen *Methanothrix* (formerly called *Methanosaeta*) (average relative abundance 1.41±0.29%) and the hydrogenotrophic methanogen *Methanobacterium* (average relative abundance 7.85±1.33%) (**Fig. A1.7**). The relative abundance of these methanogens declined during the startup period (Days 1 to 114) for the four digesters, possibly indicating that the relative abundance does not always reflect its activity (**Fig. A1.7**). It is also possible that these methanogens were replaced by another methanogen that we could not

amplify by PCR. Regardless, the decline was observed for all digesters even before monensin addition. For the four main periods of the study (P1 to P4), the average relative abundance of *Methanothrix* and *Methanobacterium* across all digester samples was 0.17±0.15% and 0.07±0.04%, respectively (**Fig. A1.7**). In the final period (P4) of the study, we observed an increase in the relative abundance of a hydrogenotrophic methanogen *Methanobrevibacter* for the Slow anaerobic digester under the high monensin concentrations, from an average of 0.09±0.04% for Days 307, 315, 321 (first three time points of P4) to a maximum of 1.25% on Day 381 (**Fig. A1.7D**). An OTU in this genus was positively correlated with monensin addition for this digester ($P<0.001$) (**Table A1.8**). It is possible that another methanogenic community member had been negatively affected by monensin addition at this concentration, because such inhibition has been described in the literature (Butsch and Bachofen 1984), however, more work is needed to substantiate this observation.

3.5 Discussion

3.5.1 Response to Low Monensin Concentrations Can Be Linked to Changes in Manure Substrate

The dairy cows excreted monensin at a rate of 13% compared to the dosing rate (**Fig. A1.1**), which is lower than previously reported for beef cattle. This resulted in relatively low monensin concentrations from 0.239±0.009 to 0.543±0.011 $mg \cdot L^{-1}$ in our manures, which is reported on a wet matter basis as $mg \cdot L^{-1}$. On a dry matter basis, the monensin concentrations in the manures ranged from 1.610±0.050 to 3.420±0.100 $mg \cdot kg^{-1}$ (**Table A1.2**). In a *prior* study with steers (beef cattle) fed with monensin at 330 $mg \cdot d^{-1}$ *per* cow, 40-50% of the parent monensin was excreted in the manure (Donoho et al. 1978). Although a study has reported on monensin concentrations in lagoon waters on two dairy farms (Watanabe et al. 2008), no studies were found in the literature

reporting monensin excretion rates for dairy cows (as opposed to beef cattle). Potentially, the difference in monensin excretion rates could be partially due to differences in dairy cow *vs.* beef cattle diets and the rates of feed passage through the animals. Regardless, the monensin concentration in the manures from our monensin-dosed cows was relatively low.

As the manure substrate was diluted to achieve the targeted organic loading rate, the Low A and Low B anaerobic digesters received manure substrate containing less than 0.2 $mg \cdot L^{-1}$ monensin. Despite this, the Low A and Low B digesters still exhibited some shifts in their anaerobic digester microbiomes as compared to the control digester (**Fig. A1.6A-B**). In addition, these anaerobic digesters showed a lower methane yield (**Fig. 3.1**). Specifically, we observed a maximum 10-15% decrease in methane yield compared to the control anaerobic digester during P3. A previous study by Thavreesi et al. (1994), also found that monensin at a similar low concentration of 0.1 $mg \cdot L^{-1}$ had decreased the performance of an upflow anaerobic sludge blanket (UASB) anaerobic digester fed with synthetic wastewater (~20% reduction in methane yield). In our study, it was necessary to establish whether the drop in performance of these anaerobic digesters was directly related to the monensin concentrations or due to some other factor.

Besides the introduction of relatively low monensin concentrations in the Low A and Low B anaerobic digesters, we observed a change in the chemical characteristics when the manure substrate was switched from the control cows to the monensin-dosed cows. The concentration of total alkalinity, total ammonium, tVFAs, as well as the pH, changed in the manures (**Table A1.2**). Together with our observation that the tVFA concentrations had not risen during the decreases in methane yields after introducing monensin-dosed manure, we believe that an indirect factor (change in manure quality) is the reason for the drop in performance observed in the Low A and Low B anaerobic digesters. In addition, the change in manure substrate can also explain the

changes in the microbiome that we observed for these anaerobic digesters. From previous research, we already knew that the quality of the substrate has an overarching effect on the anaerobic digester microbiome (Li et al. 2015b, Zhang et al. 2014). In addition, the microbiome of the monensin-dosed manure was different from the microbiome of the control manure (**Fig. A1.2A**), which may have also contributed to the change in the microbiome of the Low A and Low B anaerobic digesters (**Fig. A1.6A-B**).

3.5.2 Functional Redundancy Prevented Anaerobic Digester Failure Under High Concentrations of Monensin

The Fast and Slow anaerobic digesters both initially exhibited high levels of resistance to the 1 $mg \cdot L^{-1}$ of monensin added to the control manure substrate. The microbial community of these digesters remained relatively similar to their pre-disturbance state for this concentration of monensin (**Fig. 3.3C-D; Fig. A1.6B** (first two data points in P3 and P4, respectively)). Even though this monensin concentration was higher than our observed concentrations in the dairy cow manure, it is similar to monensin concentrations that have previously been found to be inhibitory in anaerobic digesters (Hilpert et al. 1984, Russell and Martin 1984). The immediate increase in the monensin concentration in the substrate from 1 to 5 $mg \cdot L^{-1}$ resulted in a complete loss of the functional stability for the Fast anaerobic digester, while a step-wise introduction of monensin toward the 5 $mg \cdot L^{-1}$ concentration allowed for acclimatization and functional recovery for the Slow anaerobic digester. The PCoA and taxa biplot showing only the Fast and Slow anaerobic digester samples clearly reveals that the microbiomes of these anaerobic digesters diverged under the high concentrations of monensin (**Fig. 3.4A-B**).

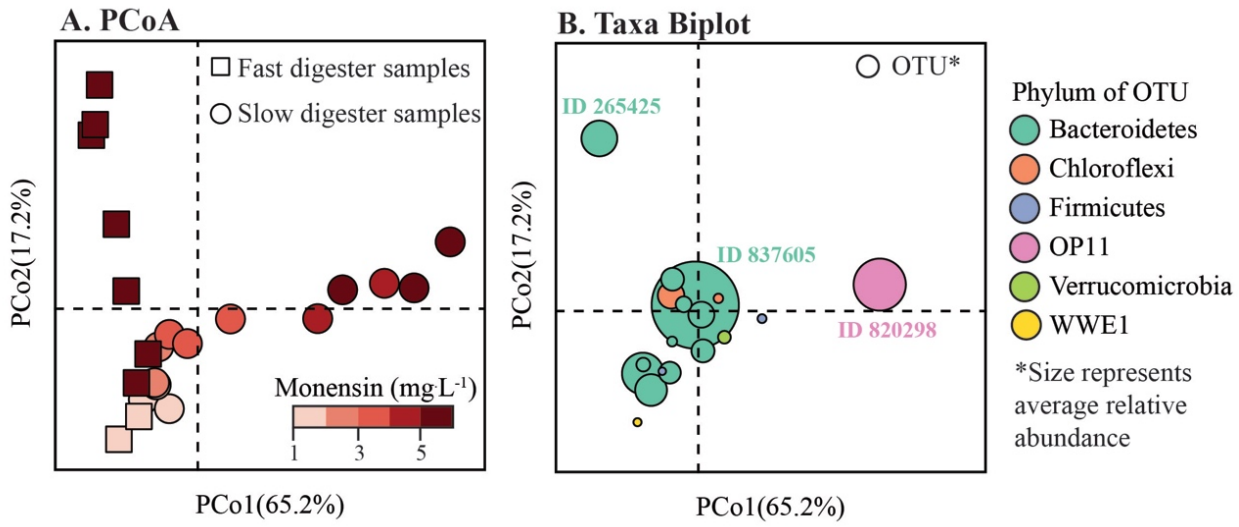

Figure 3.4. β-diversity of the Fast and Slow anaerobic digesters microbiome samples during high monensin periods (≥ 1 mg·L^{-1}; Day 265 to 307 for Fast anaerobic digester; and Day 307 to 383 for Slow anaerobic digester): **A**) principal coordinates analysis (PCoA) based on weighted Unifrac distance. Time points are colored in a gradient scale corresponding to monensin concentration in the substrate, as indicated by color sidebar; and **B**) taxa biplot showing the most abundant OTUs in these samples (OTUs that reached an average relative abundance of ≥ 1% across the time period shown). OTUs are shown nearest to the samples that they are most abundant in. Sample time points are not shown as they are displayed in the left hand PCoA. Point size represents average relative abundance of that OTU across all samples (minimum average relative abundance is 1.0%, maximum average relative abundance is 12.0%). Points are colored by phylum level taxonomy. Three OTUs discussed in the text are highlighted (OTU 820298=OP11 OTU; OTU 265425=*Prevotella* OTU; OTU 837605=Bacteriodales OTU).

In the Fast anaerobic digester, an OTU (ID 265425) belonging to the genus of *Prevotella* (order of Bacteroidales) rapidly increased in abundance, while other Bacteroidales members decreased in abundance. The biplot identified this OTU as the driver that can explain the dynamics of the microbiome (**Fig. 3.4A-B**). Other *Prevotella* spp. OTUs also increased in the Fast anaerobic digester, but to a lesser extent (data not shown). Previous studies have found varying levels of resistance to monensin within the genus of *Prevotella*, with some members of this genus requiring a period of time to adapt to this antibiotic period (Callaway and Russell 1999, Morehead and Dawson 1992, Simjee et al. 2012, Weimer et al. 2008). In the study by Callaway et al. (1999),

strains of *Prevotella bryanttii* were initially sensitive to monensin at concentrations greater than 0.7 $mg \cdot L^{-1}$. However, upon repeated exposure to monensin, certain strains were resistant to monensin at concentrations as high as 7 $mg \cdot L^{-1}$. The *Prevotella* spp. OTU relative abundance decreased in the final two days of the study (**Fig. 3.3C**), which suggests that monensin concentrations had reached an inhibitory threshold for this OTU. Because the Fast anaerobic digester did not recover its function, we cannot conclude that the *Prevotella* spp. OTU was functionally redundant with the other Bacteroidales members that had previously dominated the anaerobic digester. Different members of *Prevotella* spp. also increased in the Slow anaerobic digester upon monensin addition, though, to a much lower extent than in the Fast anaerobic digester. The specific *Prevotella* spp. OTU (ID 265425), which increased rapidly in the Fast anaerobic digester, was either not present or present in very low abundance in the Low A, Low B, and Slow anaerobic digesters throughout the study (**Figure 3.3A-B,D**).

In the Slow anaerobic digester, an OTU (ID 820298) belonging to the order d153 in the phylum of OP11 appeared to be functionally redundant with OTUs belonging to the order of Bacteroidales (**Fig. 3.3D** and **Fig. A1.3D**). When monensin was initially introduced into the digester, one OTU (ID 837605) classified to the order of Bacteroidales became dominant, reaching a maximum abundance when the monensin concentration in the substrate was 3 $mg \cdot L^{-1}$ (**Fig. 3.3D**). The subsequent decline in the Bacteroidales OTU population coincided with the rapid increase in the OP11 OTU population (**Fig. 3.3D**). During this time, the functional performance (methane yield) of the digester partially recovered (**Fig. 3.1A**). Other researchers have already reported the phylum of OP11 in anaerobic digestion, but generally as a rare species (Narihiro et al. 2015, Riviere et al. 2009). Recent work has shown that members of the OP11 phylum have genes that encode for the complete degradation of cellulose (Wrighton et al. 2014) and are likely

able to produce formate, acetate, ethanol, and lactate (Wrighton et al. 2012). A single-cell sequencing study within the phylum of OP11 showed the presence of antibiotic resistance, stress tolerance, and the possession of a Gram-negative cell wall (Youssef et al. 2011). This OTU from OP11 may have similar antibiotic resistance and stress tolerance capabilities to the strain sequenced in the previous study, which enabled it to enter the metabolic niche previously filled by Bacteriodales species. The same OP11 OTU was also increasing in the Fast anaerobic digester as well under high monensin stress conditions, though, not at the same rate, and possibly not fast enough to recover its function. This OP11 OTU appears to have been an r-strategist, which is a fast growing organism that is able to become dominant in an unstable environment (Vuono et al. 2015). It will take culturing of this OTU to conclusively state that it was functionally redundant with OTUs belonging to the order of Bacteroidales.

3.5.3 Shifts in Primary Fermentative and Syntrophic Bacterial Populations with Monensin Addition

We observed common shifts in the populations of certain primary fermenting bacteria within the order of Bacteroidales after monensin addition for all anaerobic digesters (**Fig. 3.3A-D**). Werner et al. (2011) observed in an anaerobic digester microbiome study that the primary fermenting Bacteroidetes members relied on redundancy, rather than the resiliency of individual populations, to maintain overall community function. Bacteroidales members are Gram-negative, primary fermenting bacteria, which have been shown to play important roles in polysaccharide and protein degradation, and are commonly found in anaerobic digesters (Jaenicke et al. 2011, Kampmann et al. 2012, Narihiro et al. 2015, Thomas et al. 2011). They are also generalists, which seem to be easily replaceable with other generalists (Kassen 2002).

We also found that specialists within syntrophic populations either reduced or increased their abundance during the perturbations with monensin. We do not have enough information to conclude whether these syntrophic populations showed functional resilience or redundancy. An OTU (ID 675613) belonging to the genus *Pelotomaculum*, which is a syntrophic propionate oxidizer within the order of Clostridiales (Imachi et al. 2007), was found to be positively correlated with monensin addition in all the anaerobic digesters (**Table A1.8**). On the other hand, three syntrophic OTUs (IDs 562603, 736989, and 111131) of which two belonged to the syntrophic, fatty acid degrading genus of *Syntrophus* (McInerney et al. 2007), as well as one belonging to the syntrophic, amino acid degrading genus of *Candidatus Cloacomonas* (phylum of WWE1) (Pelletier et al. 2008), were found to be negatively correlated to monensin concentrations in all anaerobic digesters ($p<0.001$) (**Table A1.9**). The genus of *Candidatus Cloacomonas* has previously been found to be sensitive to changes in anaerobic digester conditions (Kovács et al. 2013). Because the monensin concentrations were never decreased in our anaerobic digesters, it is unknown whether these OTUs would have recovered when the perturbation was lifted (functional resilience). Monensin has previously been found to decrease the precursors of methanogenesis in the cow rumen, by targeting hydrogen-, formate-, and acetate-producing bacteria and increasing propionate-producing bacteria (Chen and Wolin 1979), and thus the effect of the syntrophic bacterial populations may be pertinent.

In some instances, the increasing tVFA concentrations rather than the monensin concentration itself may have caused changes in the abundance of certain bacterial populations for the Fast or Slow anaerobic digesters. In both the Fast and Slow anaerobic digesters, Anaerolinaceae family members in the phylum Chloroflexi initially increased when the initial introduction of concentrations of monensin led to increased tVFAs concentrations (**Fig. 3.3C-D**).

A previous study by Akuzawa et al. (2011) observed an increase in a Chloroflexi subphyllum I OTU member (with 82% similarity to *Anaerolinea thermophila*) under acidogenic conditions. Some members of the Anaerolinaceae family have been reported to carbohydrate degraders that grow more rapidly when in co-culture with hydrogenotrophic methanogens (Narihiro et al. 2012, Yamada and Sekiguchi 2009).

3.6 Conclusion

Whether through direct or indirect mechanisms, monensin impacted the microbiome and performance of all the anaerobic digesters of this study. Some common shifts were observed in the primary fermenter and syntrophic populations in all digesters with monensin addition. The manure from the monensin-dosed cows that was fed to the digesters had low concentrations of monensin (less than 0.2 $mg \cdot L^{-1}$). Despite this, a shift was observed in the microbiome and performance of the anaerobic digesters that received manure from monensin-dosed cows. However, we believe that this shift was indirect and that it can be attributed to chemical differences in the monensin-dosed cow manure *vs.* the control cow manure substrates. Rapid direct introduction of higher concentrations of monensin (up to 5 $mg \cdot L^{-1}$), on the other hand, led to failure of one anaerobic digester. Slower direct introduction of the same concentrations of monensin allowed another anaerobic digester microbiome to adapt and recover its performance. The shift in the microbiome was direct from the dosing of monensin to these two digesters. In the Slow digester, an OP11 OTU appeared to be functionally redundant with members of the Bacteroidales order. Rare OTUs rapidly rose in abundance during inhibitory high monensin concentrations in both of the direct-dosed anaerobic digesters. This study builds upon previous studies in demonstrating the importance of functional redundancy in maintaining stability in anaerobic digesters. The findings of this study are based on Illumina 16S rRNA gene sequencing results. Therefore, future work

should look at how monensin and other antibiotics impact the activity of the anaerobic digester microbiome.

3.7 Acknowledgements

The authors would like to acknowledge Curt Gooch (Cornell University) for his help in formulating the study plan. The authors would like to acknowledge Dr. Julie Wittman of Elanco Animal Health for her support and advice during the cow and anaerobic digester portion of the study. The authors would like to thank Gladys Birdsall (Cornell University Teaching and Research Barn) for her assistance with the cow portion of the study and Joseph Usack (Cornell University) for providing training and assistance in anaerobic digester operation and maintenance. Finally, the authors would like to acknowledge members of the Knight Lab (University of California, San Diego), specifically Antonio Gonzalez Pena, Yoshiki Vazquez Baezam, and Gail Ackermann, for their assistance with sequencing of the cow hindgut samples, as well as advice on sequencing data analysis.

CHAPTER 4

HIGHER SUBSTRATE RATIOS OF ETHANOL TO ACETATE DRIVE PRODUCTION TOWARD *N*-CAPRYLATE IN BIOREACTOR WITH PRODUCT EXTRACTION

Adapted from: Spirito, Marzilli, and Angenent. In preparation for *Environmental Science and Technology*.

4.1 Abstract

Anaerobic reactor microbiomes are capable of converting organic wastes to medium-chain carboxylates *via* a process termed reverse beta-oxidation. Generally, an electron-rich substrate, such as one containing ethanol or lactic acid, is used to drive the chain elongation of shorter-chain carboxylates, such as acetate and *n*-butyrate, to longer-chain carboxylates, such as *n*-caproate or *n*-caprylate. Recent research has looked at coupling syngas fermentation, which produces principally ethanol and some acetate, to chain-elongation bioreactors to upgrade the produced ethanol to *n*-caproate or *n*-caprylate. To do this, it is important to understand what ratios of ethanol-to-acetate produced by syngas fermentation lead to optimal medium-chain carboxylate production. Here, we examined the effect of different ethanol-to-acetate substrate ratios on the production of medium-chain carboxylates in a 0.7-L upflow anaerobic filter bioreactor with product extraction. Over an approximately eight-month period, we characterized the performance and the microbiome composition (*via* Illumina 16S rRNA gene sequencing) of the chain-elongation bioreactor. Higher ethanol-to-acetate substrate ratios led to higher selectivity for *n*-caprylate in the products of the bioreactor. The highest *n*-caprylate selectivity in this study occurred when the substrate contained primarily ethanol (ethanol-to-acetate ratio > 200, on a COD basis), however, a large fraction of the

fed ethanol was not consumed and was wasted in the effluent. At an approximately one-to-one substrate COD ratio, *n*-caprylate production stopped in the bioreactor. The experimental results in this study generally backed up the predictions of a generalized thermodynamic model that we developed to predict the thermodynamic feasibility of *n*-caprylate production at different ethanol-to-acetate ratios. The bioreactor microbiome in this study was found to be relatively uneven and dominated by a few OTUs primarily belonging to the Firmicutes and Proteobacteria phyla. In conclusion, this study indicates that a syngas fermentation effluent that has higher ethanol-to-acetate concentrations should lead to higher *n*-caprylate productivity when fed to a chain-elongation bioreactor, however, care should be taken not to overload the chain-elongation bioreactor system.

4.2 Introduction

In syngas fermentation, pyrolysis gas (a mixture of CO, CO_2, and H_2 gas) from the gasification of organic waste materials (*e.g.,* lignocellulosic biomass, municipal solid waste, industrial waste, or agricultural waste), as well as off-gas from steel manufacturing, can be converted into ethanol by anaerobic gas-fermenting bacteria (*i.e.,* acetogens) (Daniell et al. 2012, Liew et al. 2016, Molitor et al. 2016, Munasinghe and Khanal 2010). Ethanol production rates of higher than 150 g L^{-1} day^{-1} have been reported in syngas fermentation systems (Gaddy et al. 2007, Molitor et al. 2016) and several commercial syngas fermentation facilities are being built (Liew et al. 2016, Molitor et al. 2016). The concentration of ethanol produced by syngas fermentation is typically limited to 1-6 wt%, due to the toxicity of the ethanol product to the bacteria involved (Molitor et al. 2016). Extraction of the dilute ethanol from solution *via* distillation is energetically costly. For example, a paper by Vane et al. (2008) estimates that for fractional distillation of 6 wt% ethanol solutions,

5 MJ-fuel is required per kg of ethanol distilled.

Recently, researchers have investigated coupling the syngas fermentation to chain elongation to upgrade the dilute ethanol and produce medium-chain carboxylates in either one-stage (Diender et al. 2016, Richter et al. 2016) or two-stage systems (Gildemyn et al. 2017, Vasudevan et al. 2014) (Kucek et al. 2016c). The chain-elongation process is a modification of the anaerobic digestion process. However, instead of methane, the anaerobic reactor microbiome is directed to produce medium-chain carboxylates, such as *n*-caproate, *n*-heptanoate, and *n*-caprylate. Specifically, an electron donor, such as ethanol or lactic acid, is used to drive the chain elongation of short-chain carboxylates, such as acetate and *n*-butyrate, to produce medium-chain carboxylates (Angenent et al. 2016, Cavalcante et al. 2017, Spirito et al. 2014). To limit methane production in these systems, one strategy that has been employed involves lowering the pH in the bioreactor to ~5-5.5 and using membrane-based liquid-liquid extraction (pertraction) to remove the relatively hydrophobic medium-chain carboxylate products (Agler et al. 2012a, Kucek et al. 2016c). Medium-chain carboxylates can potentially be used industrially as antimicrobials in animal feed (Desbois 2012), as fragrances and flavors (Kenealy et al. 1995, Tao et al. 2014), or upgraded to diesel fuel (Levy et al. 1984, Levy et al. 1981) or aviation fuel (Harvey and Meylemans 2014).

The ratio of ethanol to acetate produced in syngas fermentation is typically high, with reports of ratios of ~30 (based on a COD) in optimized systems (Phillips et al. 1993, Richter et al. 2013). One recent paper by Abubackar et al. (2016) has even reported the production of mainly ethanol *via* syngas fermentation by *Clostridium autoethanogenum* DSM 100061. However, the rates were relatively low. To optimize the use of syngas fermentation effluent in the chain-elongation process, it is necessary to determine the optimal ratio of ethanol to acetate needed in syngas fermentation effluent to direct the chain-elongation process toward production of longer

chain carboxylates, such as *n*-caprylate. Previous research has shown that higher ratios of ethanol to acetate lead to higher ratios of *n*-caproate to *n*-butyrate. Pure culture studies with *Clostridium kluyveri* have demonstrated this in both batch (Bornstein and Barker 1948, Weimer and Stevenson 2012) and continuously fed (Kenealy and Waselefsky 1985) bioreactors. Similarly, mixed culture studies have found that higher ethanol concentrations can drive chain elongation toward *n*-caproate, as long as the ethanol concentrations do not become inhibitory (Liu et al. 2016, Lonkar et al. 2016, Weimer et al. 2015). Angenent et al. (2016) recently developed a simple generalized stoichiometric model to predict the thermodynamic favorability of *n*-caproate formation at different ethanol-to-acetate substrate ratios. The model, which did not extend to *n*-caprylate formation, demonstrated that increased ethanol-to-acetate ratios in the substrate created more thermodynamically favorable conditions for chain elongation to *n*-caproate. At higher ethanol-to-acetate ratios consumed, more free energy is released by the fermentation, which makes ATP production and the complete chain-elongation process more thermodynamically feasible. A recent study by Kucek et al. (2016c) used batch tests to demonstrate that increasing concentrations of ethanol to acetate also lead to increased production of *n*-caprylate compared to *n*-caproate. In addition, Kucek et al. (2016c) tested two different substrate ratios of ethanol to acetate (6 and 15, on a COD basis) in continuous anaerobic filter bioreactor and found higher *n*-caprylate-to-*n*-caproate productivity at the higher ratio.

Here, our main objective was to experimentally determine what substrate ethanol-to-acetate ratio was optimal to promote *n*-caprylate production in a continuously operated bioreactor with product extraction. In addition, this study aimed to investigate whether *n*-caprylate could be produced from primarily ethanol in the substrate (*i.e.,* ethanol-to-acetate substrate COD ratio >200). The thermodynamic model of Angenent et al. (2016) was extended in this study to *n*-

caprylate and the experimental data collected was used to validate the model. Finally, this study used Illumina 16S rRNA gene sequencing to investigate what OTUs were present in the bioreactor microbiome during the operating period and correlated the relative abundance of these OTUs with *n*-caprylate specificities.

4.3 Materials and Methods

4.3.1 Bioreactor Setup

The bioreactor system that was operated in this study was previously used and extensively described in the study by Kucek et al. (2016c). Briefly, an upflow anaerobic filter (working volume 0.7 L) was operated with a continuous in-line, membrane-based liquid-liquid extraction (*i.e.,* pertraction) system. The feed rate used in this study was approximately 0.6 L d^{-1}, while the system recycle flow rate was 130 L d^{-1}, which resulted in a recycle feed ratio of higher than 200. The hydraulic retention time (HRT) that was used in this study was ~1.2 days. The temperature of the bioreactor was maintained at 30±1ºC. The pH of the bioreactor broth was maintained at 5.26±0.09 *via* addition of 0.5 M hydrochloric acid to the well-mixed feed and recycle inlet at the base of the bioreactor. The pH of the alkaline extraction solution was maintained at 9.48±0.34 *via* addition of 5 M sodium hydroxide.

4.3.2 Growth Medium and Inoculum

The growth medium used in this study has been described previously (Kucek et al. 2016c, Vasudevan et al. 2014). For each operating period, the substrate concentrations of ethanol and acetate in the growth medium were varied to achieve the targeted substrate ratios of ethanol and acetate, while maintaining organic loading rates (OLR) in the range of 18.7 to 28.2 g COD L^{-1} d^{-1}

(**Table 4.1**). As mentioned above, this bioreactor microbiome was also used in a previous study by Kucek et al. (2016c). At the end of the Kucek et al. (2016c) study, the bioreactor was overloaded and its performance suffered. Between the end of the Kucek et al. (2016c) study and our current study, a period of approximately six months passed during which bioreactor performance was recovered (data not shown). No new inoculum was added to the bioreactor between studies.

Table 4.1. Operating data for the bioreactor. Average hydraulic retention times (HRT), influent ethanol and acetate concentration, substrate COD ratios (ethanol-to-acetate), and organic loading rates (OLR) *per* period are reported as mean ± s.e. During Period 9, the extraction system was off for the bioreactor. *For Period 10, measurements of acetate were not made in the influent so values reported are approximate. Data *prior* to Period 1 not included in this chapter.

Period	**Days**	**HRT**	**Substrate COD Ratio**	**Ethanol (mM)**	**Acetate (mM)**	**OLR (g COD L^{-1} d^{-1})**
Period 1	155 to 197	1.2 ± 0.02	11.7 ± 1.0	200.42 ± 9.42	25.68 ± 1.75	18.72 ± 0.81
Period 2	198 to 210	1.17 ± 0.05	274.9 ± 58.3	289.14 ± 3.64	1.58 ± 0.33	25.15 ± 1.1
Period 3	211 to 226	1.27 ± 0.07	16.9 ± 2.1	281.97 ± 5.21	24.97 ± 3.01	23.88 ± 1.43
Period 4	227 to 239	1.16 ± 0.04	6.7 ± 0.4	281.69 ± 10.5	63.35 ± 3.35	28.17 ± 1.24
Period 5	240 to 256	1.17 ± 0.04	3.6 ± 0.3	228.81 ± 10.65	94.09 ± 4.74	25.19 ± 1.18
Period 6	257 to 274	1.2 ± 0.05	2.9 ± 0.1	216.13 ± 5.26	111.83 ± 3.41	24.63 ± 1.08
Period 7	275 to 290	1.2 ± 0.04	1.8 ± 0.1	177.87 ± 4.53	146.25 ± 6.24	23.34 ± 0.85
Period 8	291 to 336	1.14 ± 0.04	5.1 ± 0.5	218.15 ± 9.68	64.64 ± 6.1	23.46 ± 1.16
Period 9	337 to 357	1.2 ± 0.04	6.5 ± 0.5	245.15 ± 11.92	56.39 ± 3.55	24 ± 1.31
Period 10	358 to 369	1.14 ± 0.04	6*	220.82 ± 17.14	NA	24*
Period 11	370 to 397	1.12 ± 0.05	135.2 ± 18.1	286.51 ± 6.08	3.18 ± 0.42	26.22 ± 1.35

4.3.3 Bioreactor Operation

In the main phase of this study (Periods 1 to 7), we operated the bioreactor at the following substrate (ethanol-to-acetate) COD ratios: 1.8, 2.9, 3.6, 6.7, 16.9, 11.7, and 274.9 (**Table 4.1**). We operated the bioreactor at each ratio for a period of at least two weeks (at least 11 HRTs). Following a loss of performance at the 1.8 COD ratio, we ran another set of similar substrate COD ratios (Periods 8 to 11): 5.1, 6.5, 6.0, and 135.2. The HRTs, OLRs, and additional details about the periods in the study can be found in **Table 4.1**.

4.3.4 Liquid and Gas Analysis

Liquid samples (1.5 mL) were collected from the bioreactor influent, the bioreactor broth, and the alkaline extraction solution, as has been described by Kucek et al. (2016c). Analysis of the samples to determine carboxylate and ethanol concentrations was carried out by gas chromatography using the method outline by Usack et al. (2015). The concentrations of methane, carbon dioxide, and hydrogen gases (detection limit 0.2%) were measured using a GC system described in Usack et al. (2015).

4.3.5 Calculations and Statistical Analysis of Operating Data

Carboxylate production rates were calculated as the average values for each operating period. Average effluent production rates per liter of bioreactor (g COD L^{-1} d^{-1}) and average transfer rates *via* product extraction (g COD L^{-1} d^{-1}) were summed to yield the total production rates per liter of bioreactor (g COD L^{-1} d^{-1}). COD stands for chemical oxygen demand. The average effluent production rates were calculated by dividing the average carboxylate concentration *per* period by the average HRT for that period. The average HRT *per* period was calculated based on the average

influent flow rate *per* period, which was determined volumetrically. The average transfer rates were calculated by first plotting the increasing concentrations of the individual carboxylates in the alkaline extraction solution *vs.* time. Then the linear model function, lm, in the stats package in R v.3.3.2 (R Core Team 2016) was used to determine the slope and standard error of the best-fit line through these points. The slope was then divided by the bioreactor working volume (0.7 L) to get the average transfer rate *per* period. We calculated the conversion efficiency as the individual carboxylate total production rate divided by the organic loading rate *per* period. In addition, specificity was calculated as the individual carboxylate total production rate divided by the combined total production rate for all carboxylates during each period (where the carboxylates included are *n*-butyrate, *n*-caproate, and *n*-caprylate). Furthermore, pertraction efficiency was calculated as the average transfer rate divided by the total production rate for each carboxylate. RStudio v1.0.136 (RStudio Team 2016) was used to perform all data analyses in R. All concentrations, rates, ratios, and efficiencies are reported as mean value ± standard error in the paper, unless indicated otherwise in the text.

4.3.6 Thermodynamic Model Development

Here, we extended the generalized stoichiometric model developed by Angenent et al. (2016) to predict the thermodynamic favorability of *n*-caprylate formation at different substrate ratios (**Figure 4.1**). Briefly, the model uses stoichiometric relationships to predict the moles of caprylate and ATP that would be produced based on the moles of ethanol and acetate provided to the system. For the purpose of this model, the moles of *n*-butyrate and *n*-caproate formed were also fixed to variables “b” and “c”, respectively. The boundary for the metabolic flux was set to 10 moles of ethanol and acetate combined (therefore by setting moles of ethanol is equal to “a” in the model,

moles of acetate is set to "10-a"). This resulted in the overall stoichiometric equation:

$aC_2H_6O + (10\text{-}a)\ C_2H_3O_2^- + (10\text{-}a)\ H^+ \rightarrow b\ C_4H_7O_2^- + c\ C_6H_{11}O_2^- + (10\text{-}2b\text{-}3c)/4\ C_8H_{15}O_2^- + (30\text{-}2a\text{-}2b\text{-}c)/2\ H_2O + (4a+2b+c\text{-}30)/2\ H_2 + (10+2b+c)/4\ H^+ + (30\text{-}2b\text{-}c)/8$ ATP

In our model, the stoichiometry of all metabolites (*i.e.,* ethanol, acetate, *n*-butyrate, *n*-caproate, and *n*-caprylate, molecular hydrogen, water, intermediary metabolites, redox mediators, and ATP) depends on the variables "a", "b", and "c". The stoichiometry for re-oxidation of reduced ferredoxin *via* H2-ase or Rnf and ATP synthase varies depending on the variables "a", "b", and "c", which determines the molecular hydrogen production and ATP production, respectively (**Figure 4.1**). The net consumption of one mole of water during acetate production from acetyl-CoA *via* substrate level phosphorylation (due to ATP hydrolysis) was considered when balancing the overall equation. Based on the overall stoichiometry and the ethanol and carboxylate concentrations that were measured in the bioreactor, the Gibbs free energy of the reaction was calculated, as well as the Gibbs free energy required for ATP formation. It was assumed that -72 kJ was required *per* mole of ATP produced, though we note that this value would vary based on the concentration of inorganic phosphate present in the bioreactor (Buckel and Thauer 2013). If the Gibbs free energy of the reaction was more negative than the Gibbs free energy required for ATP formation, the reaction was deemed feasible. For the purpose of these calculations, the standard Gibbs free energy of formation values for the reactants and products were taken from Kleerebezem and van Loosdrecht (2010) with the exception of *n*-caprylate, which was not available. The standard Gibbs free energy of formation for *n*-caprylate was calculated to be -323.8 kJ mol^{-1} using the group contribution method described by Mavrovouniotis (1990). In our model, we established three different scenarios for the balance of *n*-butyrate, *n*-caproate, and *n*-caprylate that were produced at the different ethanol-to-acetate ratios experienced by the microbiome. For

Scenario 1, we assumed only *n*-caprylate was formed by the microbiome. For Scenario 2, we assumed 1 mole of *n*-butyrate, 1 mole of *n*-caproate, and 1.25 moles of *n*-caprylate were formed by the microbiome. Finally, for Scenario 3, we assumed 1.5 moles of *n*-butyrate, 1.5 moles of *n*-caproate, and 0.625 moles of *n*-caprylate were formed (**Fig. A2.1**).

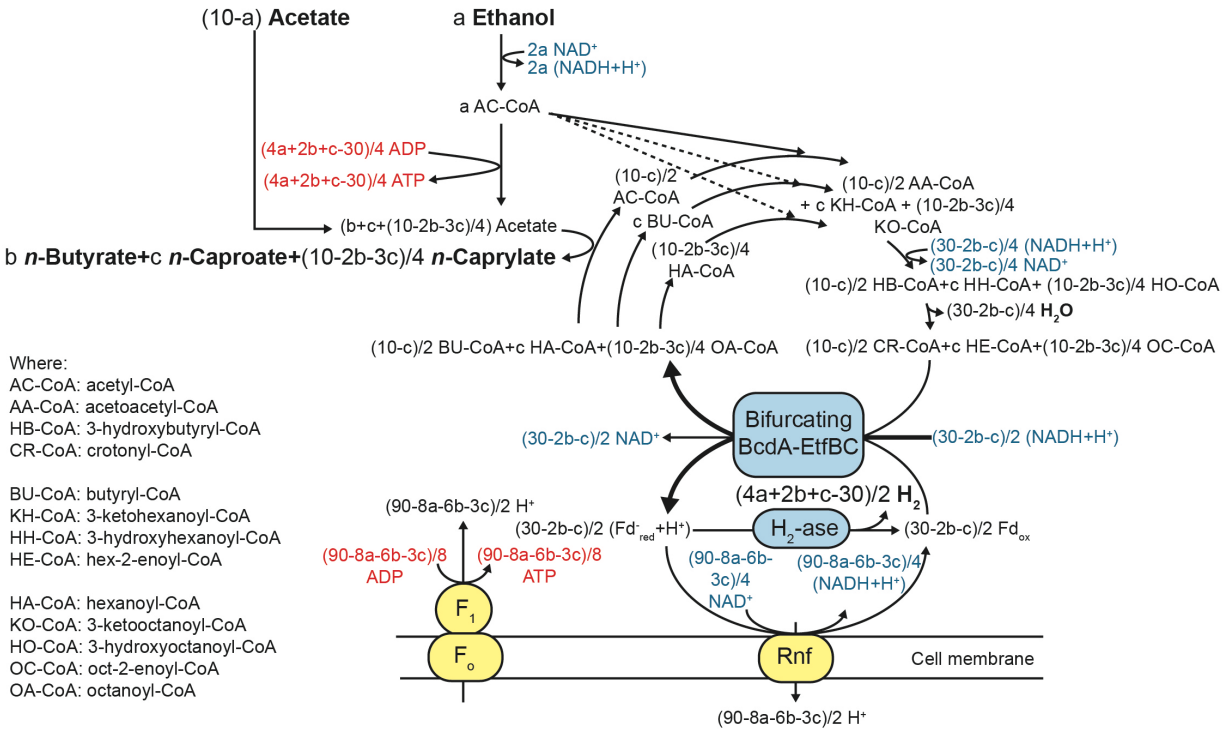

Figure 4.1. Generalized stoichiometric model for the fermentation of ethanol and acetate to *n*-butyrate, *n*-caproate, *n*-caprylate, and molecular hydrogen by *C. kluyveri*. This model is the extended version of the model developed by Angenent et al. (2016). The variable "a" represents moles of ethanol, "b" represents moles of *n*-butyrate, "c" represents moles of *n*-caproate. Redox factors are highlighted in blue; classical energy conservation in red; and more recently described mechanisms of energy conservation in yellow. F_0/F_1 is H^+/Na^+-pumping ATP synthase complex and Rnf is the ferrodoxin-NAD reductase complex.

4.3.7 Microbial Community Analysis

We collected biomass samples for Illumina 16S rRNA gene sequencing analysis from the bottom

and middle of the bioreactor approximately weekly throughout the operating period. The method of collecting biomass samples from the bioreactor broth has been previously described by Kucek et al. (2016c). However, in the previous study, biomass samples were only collected from the middle of the bioreactor, whereas in this study we also collected them from a bottom sampling port. Pelleted biomass samples were stored at -80°C until further processing.

Genomic DNA was extracted using the PowerSoil-htp 96 Well Soil DNA Isolation kit (MO BIO Laboratories Inc., Carlsbad, CA) according to the protocol of the manufacturer. The DNA amplification protocol was described previously (Regueiro et al. 2015) with the following exceptions: Mag-Bind RxnPure Plus magnetic beads solution (Omega Biotek, Norcross, GA, USA) was used instead of Mag-Bind E-Z Pure and 50 ng DNA *per* sample was pooled instead of 100 ng. Duplicate PCR reactions of each DNA extract were performed and pooled *prior* to sequencing. Paired-end reads were joined in QIIME version 1.9.1 (Caporaso et al. 2010) using the joined_paired_ends.py script and then the joined reads were uploaded to QIITA (qiita.microbio.me) for further processing. The sortmerna method (Kopylova et al. 2012) was used to bin sequences in operational taxonomic units (OTUs) at 97% identity. Taxonomy was assigned for representative sequences selected for each OTU using the Greengenes v13.8 database from August 2013 (McDonald et al. 2012). The remaining analyses were performed locally in QIIME v1.9.1 and RStudio v1.0.136 (RStudio Team 2016). Singletons were removed from the dataset resulting in 932 unique OTUs.

Alpha diversity was analyzed *via* the Gini coefficient (a measure of unevenness), observed OTUs (richness), and Shannon diversity (Shannon 1948) metrics available in QIIME v1.9.1. 100 rarefactions at a depth of 6510 sequences per sample were performed and collated. Statistical analyses of the alpha diversity results were performed using analysis of variance (ANOVA) and

Tukey HSD in R. In addition, heat maps were created to represent OTU relative abundance *via* the gplots package in R (Warnes et al. 2015). Correlations of OTU relative abundance with *n*-caprylate specificities was investigated using the Spearman's rank correlation coefficient *via* rcorr function (Hollander and Wolfe 1973, Press et al. 1988) of the Hmisc package in R v.3.3.2. Correlations with $p<0.001$ were considered significant. Only OTUs that reached at least 1% relative abundance in any one bioreactor sample were considered in the correlation analysis.

Sequences were submitted to EBI under the accession number ERP024133. Sequences and study metadata are publically available in QIITA, which is an open-source microbiome storage and analysis resource, under study number 11161.

4.4 Results and Discussion

4.4.1 Higher Ethanol-to-Acetate Substrate Ratio Leads to Higher N-Caprylate Specificity

We observed that higher ethanol-to-acetate ratios in the substrate led to higher *n*-caprylate specificities in the products of the bioreactor (**Fig. 4.2, Table A2.1**). The highest *n*-caprylate specificity achieved in this study was 82 ± 10% at the highest substrate ratio tested, 274.9 ± 58.3 (**Table A2.1**). To our knowledge, this is the first time that only ethanol without much acetate was used as a substrate for medium-chain carboxylate production. Generally, as the ethanol-to-acetate substrate ratio was lowered, the *n*-caprylate specificity decreased (**Fig. 4.2**). At the lowest substrate ratio used in this study (*i.e.,* 1.8 ± 0.1 g COD of ethanol *per* g COD of acetate), no *n*-caprylate production was observed and *n*-butyrate specificity was much higher than at the higher substrate ratios tested (**Fig. 4.2, Table A2.1**).

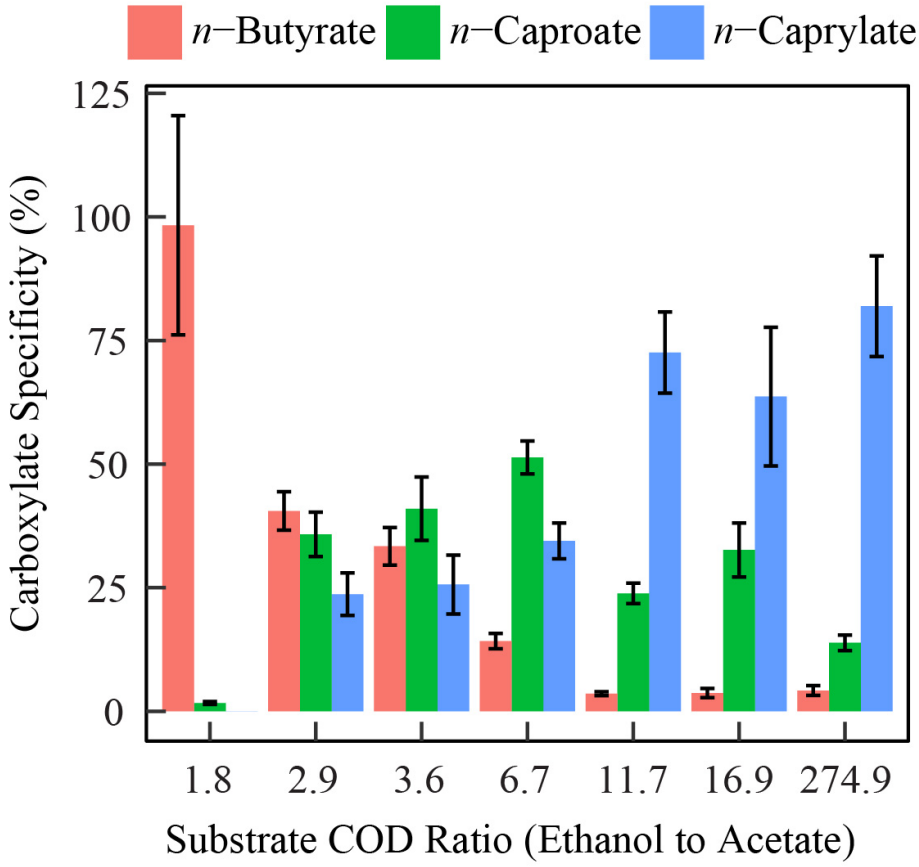

Figure 4.2. Specificity of carboxylate produced at each substrate ratio during the main periods of the study (Periods 1 to 7) for *n*-butyrate, *n*-caproate, and *n*-caprylate. Error bars indicate standard error.

With the exception of Period 1 where the ethanol-to-acetate substrate ratio was 11.7 ± 1.0, we decreased the substrate ratio during the main operating period from 274.9 ± 58.3 in Period 2 to 1.8 ± 0.1 in Period 7 (**Table 4.1**). When the substrate ratio of ethanol to acetate decreased, acetate and *n*-butyrate concentrations increased in the effluent of the bioreactor (**Fig. 4.3A, Table A2.2)**. In Period 2, at the highest substrate ratio (274.9 ± 58.3) used in this study, the average acetate and *n*-butyrate concentrations measured in the effluent leaving the bioreactor were 4.1 ± 0.6 mM and 3.6 ± 0.9 mM, respectively (**Fig. 4.3A, Table A2.2**). As the substrate ratio was decreased

throughout the course of the study (Periods 2 to 7), the acetate and *n*-butyrate concentrations in the bioreactor increased. At the lowest substrate ratio employed in this study (1.8 ± 0.1) in Period 7, the average acetate and *n*-butyrate concentrations in the effluent were 88.2 ± 8.4 mM and 18.3 ± 0.9 mM, respectively (**Fig. 4.3A, Table A2.2**). Since the bioreactor had a high recycle ratio of ~220, the concentration leaving in the effluent can be considered approximately equal to the concentration in the bioreactor, though, slight differences may exist, as will be discussed later.

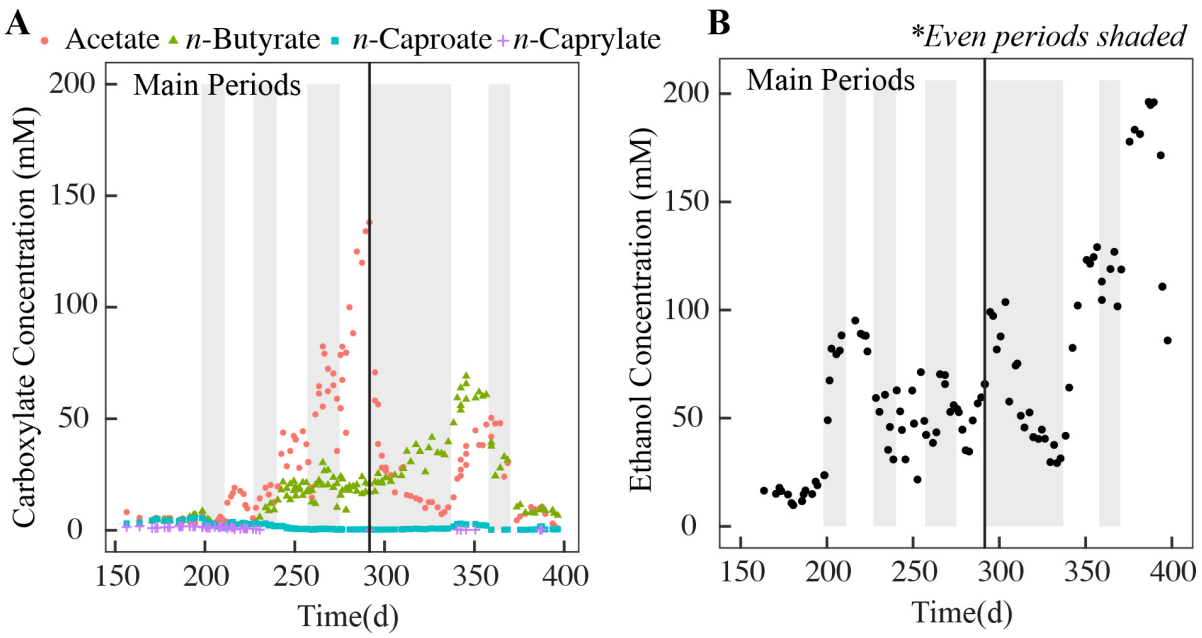

Figure 4.3. Effluent concentrations of carboxylates (A) and ethanol (B) during the operating period in the bioreactor. Shaded sections represent even periods (*i.e.,* Period 2, Period 4, Period 6, Period 8, Period 10). Main periods (Periods 1 to 7) in study are before the solid vertical black lines in plots. Data *prior* to Period 1 not shown.

We extended the thermodynamic model originally developed by Angenent et al. (2016) to predict the thermodynamic favorability of *n*-caprylate formation at the different ethanol-to-acetate ratios experienced by the bioreactor microbiome. Since the bioreactor was well-mixed, we used

the average measured effluent concentrations (**Fig. 4.3A and B, Table A2.2**) in this model to represent the closest approximation of the conditions the microbiome saw. For the most part, our model described what we observed experimentally. Similar to the trend observed by Angenent et al. (2016) for *n*-caproate, our model predicted that higher ethanol-to-acetate ratios experienced by the microbiome led to more favorable thermodynamic conditions for chain elongation to *n*-caprylate (**Fig. A2.1).** In other words, the Gibbs free energy of the *n*-caprylate formation reaction was more negative than the Gibbs free energy required for ATP production at higher ethanol-to-acetate ratios. At the higher substrate ratios employed in Periods 2 and 3 (*i.e.,* 274.9 and 16.9, on a COD basis; **Table 4.1**), the measured molar ethanol-to-acetate ratio in the bioreactor effluent was relatively high (*i.e.*, 16.5 and 7.0, respectively; **Table A2.2**). At these ratios, our model predicted that formation of only *n*-caprylate was thermodynamically feasible (*i.e.,* $|\Delta G_{reaction}|>|\Delta G_{required}|$ for scenario 1 in the model; **Fig. A2.1**). Indeed, we observed high specificities (over 60%) for *n*-caprylate formation in our bioreactor during Periods 2 and 3 (**Fig. 4.2 and Table A2.1**). In Period 4, at a substrate COD ratio of 6.7 (**Table 4.1**), the measured molar ethanol-to-acetate ratio in the bioreactor effluent declined to 3.1 (**Table A2.2**). At this ratio observed by the microbiome, the model predicted that formation of only *n*-caprylate was not thermodynamically feasible, but formation of a mix of *n*-butyrate, *n*-caproate, and *n*-caprylate was (*i.e.,* $|\Delta G_{reaction}|>|\Delta G_{required}|$ for scenarios 2 and 3 but not scenario 1 in the model; **Fig. A2.1**). During Period 4, the product spectrum did become more mixed, with *n*-caproate specificity exceeding *n*-caprylate specificity in the products of the bioreactor and *n*-butyrate specificity also increasing (**Fig. 4.2 and Table A2.1**). At the lowest substrate COD ratios employed in the main periods of the study (*i.e.*, 3.6, 2.9 and 1.8 in Periods 5, 6, and 7, respectively, **Table 4.1**), the measured ethanol-to-acetate molar ratios in the effluent of the bioreactor were also correspondingly low (*i.e.,*

1.4, 1.0, and 0.6, respectively, **Table A2.2**). At these ratios, the model predicted that *n*-caprylate formation was not thermodynamically feasible at any of the scenarios tested (**Fig. A2.1**). Nevertheless, we did observe some *n*-caprylate production in Periods 5 and 6 (**Fig. 4.2**). The model also predicted hydrogen consumption at these low ratios (not shown), which was likely not the reality in the bioreactor. Finally, in Period 1, the conversion efficiency of ethanol and acetate to medium-chain carboxylates was high (**Table A2.3**) and ethanol and acetate concentrations in the effluent of the bioreactor were correspondingly low (**Table A2.2**). Based on the effluent concentrations measured in this period, the model incorrectly predicted that *n*-caprylate formation was not feasible in the bioreactor (**Fig. A2.1**). Our model is an over-simplification of what occurred in the bioreactor and does not fully represent the environment experienced by the microbes. In addition, our thermodynamic model does not account for the kinetics in the bioreactor system. Finally, the model does not account for effect of the continuously operating pertraction system, which was continuously and preferentially removing longer chain carboxylates from the bioreactor environment.

4.4.2 N-Caprylate Production with Primarily Ethanol in Substrate but Lower Overall Productivity

The highest substrate ratio resulted in the highest *n*-caprylate specificity (**Fig. 4.2**), as well as the highest *n*-caprylate-to-*n*-caproate productivity ratio of 5.9 ± 0.7 (**Table A2.1**). However, with the increase in substrate ratio from Periods 1 to 2 (*i.e.,* from 274.9 to 11.7), there was a decrease in the substrate conversion efficiency to *n*-caprylate (*i.e.,* from 68 ± 7% in Period 1 to 43 ± 5% in Period 2; see **Table A2.3**). In a recent pure culture study with *Clostridium kluyveri* by Gildemyn et al. (2017), the authors also observed that a higher ethanol-to-acetate substrate ratio (10 *vs.* 3) led to a

higher *n*-caprylate specificity but a decreased carbon conversion efficiency to *n*-butyrate, *n*-caproate, and *n*-caprylate. Similarly, the study by Kucek et al. (2016c) found that their highest *n*-caprylate specificity was not associated with their highest medium-chain conversion efficiency. It appears that at higher ethanol-to-acetate ratios there is a trade-off between improved *n*-caprylate specificity and decreased overall productivity.

In Period 1, the average ethanol concentration in the effluent of the bioreactor was 15.6 ± 1 mM (**Table A2.2**). Under the higher ethanol-to-acetate ratio in Period 2, the ethanol concentration increased in the effluent of the bioreactor and reached a maximum of 95 mM on day 216 near the beginning of Period 3 (**Fig. 4.3B)**. The concentration remained well below the concentrations that have previously been found to be inhibitory in a pure-culture study (460 mM) (Weimer and Stevenson 2012) and in a mixed-culture study (~300 mM) (Kucek et al. 2016c). However, it is clear that excess ethanol was not being fully utilized by the bioreactor. During periods of high medium-chain conversion efficiency (greater than 70%) in the study by Kucek et al. (2016c), ethanol concentrations remained below ~10 mM in their bioreactor effluent. Following period 2 in our study, ethanol concentrations never returned to the low concentrations seen in Period 1 (**Fig. 4.3B, Table A2.2**), despite lowering of the substrate ethanol-to-acetate ratio (**Table 1**). In Periods 3 to 7, the production rates of *n*-caprylate decreased (**Table A2.4**) and the medium-chain conversion efficiency remained below 60% (**Table A2.3**). We employed higher organic loading rates (*i.e.,* 18.7 to 28.2 g COD L^{-1} d^{-1}, see **Table 1**) than the study by Kucek et al. (2016c), where the organic loading rates were 6.3, 13.7, and 15.0 g COD L^{-1} d^{-1} during the periods with the highest medium-chain conversion efficiencies (*i.e.,* greater than 70%). Lowering the overall organic loading rate may have increased performance (*i.e.,* conversion efficiency of the substrate to *n*-caprylate) in our study, by preventing the accumulation of ethanol in the bioreactor.

In our study, relatively short HRTs (*i.e.,* 1.2 days) were employed. A longer HRT may have allowed for more of the *n*-butyrate produced to be converted to *n*-caproate or *n*-caprylate, which could have resulted in improved production at these higher ethanol-to-acetate substrate ratios.

n-Caproate and *n*-caprylate contain six or eight carbon atoms in their chain, respectively. Thus, they are relatively hydrophobic and easy to extract from solution, compared to *n*-butyrate. Indeed, it can be seen that the average extraction efficiency *per* period for *n*-butyrate ranged from a minimum of 11.8% to a maximum of 66.8%, whereas the average extraction efficiency for *n*-caproate *per* period was always greater than 65% and for *n*-caprylate was always greater than 96% in the periods when *n*-caprylate production was observed (**Table A2.5**). Due to the high extraction efficiency for *n*-caprylate that we observed in this study, it is unlikely that *n*-caprylate production rates were limited by mass transfer limitations.

4.4.3 Microbiome Shifts Correlated to N-Caprylate Specificity

We performed a time-series analysis of the microbiome that was sampled from two locations in our bioreactor: a bottom and a middle sampling port. We used Illumina 16S rRNA gene sequencing to analyze the samples. Similar to the microbial community stratification that has previously been observed in upflow anaerobic sludge blanket reactors (MacLeod et al. 1990), we observed clear differences in the compositions of the microbiomes sampled from the bottom and the middle of the bioreactor. The majority of samples from the bottom of the bioreactor had a higher relative abundance of the phylum Firmicutes compared to the phylum Proteobacteria (**Fig. A2.2A**), whereas the reverse was true in the middle of the bioreactor (**Fig. A2.2B**). In addition, the bottom of the bioreactor had a significantly more diverse microbiome, as indicated by the Shannon

diversity index (4.634 ± 0.090 for the bottom *vs*. 3.973 ± 0.054 for the middle port; p<<0.0001), and a less uneven microbiome, as indicated by the Gini coefficient (0.984 ± 0.000 for the bottom *vs.* 0.987 ± 0.000 for the middle port; p<<0.0001; note that a lower value reflects a more even community) (**Table A2.6**). The Shannon diversity index accounts for both the richness and evenness of the microbial community structure (Shannon 1948). Though the bioreactor was well-mixed due to the high recycle rate employed, it is possible that slight differences existed throughout the bioreactor. Substrate and acid (0.5M HCl) for pH control were added at the base of the bioreactor, which may have caused slightly higher concentrations of un-dissociated carboxylic acids and ethanol to be seen by the microbes at the bottom of the bioreactor, as compared to the middle. Nevertheless, we did observe common OTUs between the middle and the bottom bioreactor samples, which are indicated on the heat maps (**Fig. 4.4 and 4.5**). Of the OTUs that reached higher than one percent relative abundance in the bioreactor samples from the bottom and middle of the bioreactor (40 and 45 OTUs, respectively) and are shown in the heat maps, 28 of these OTUs were shared between the two sampling locations.

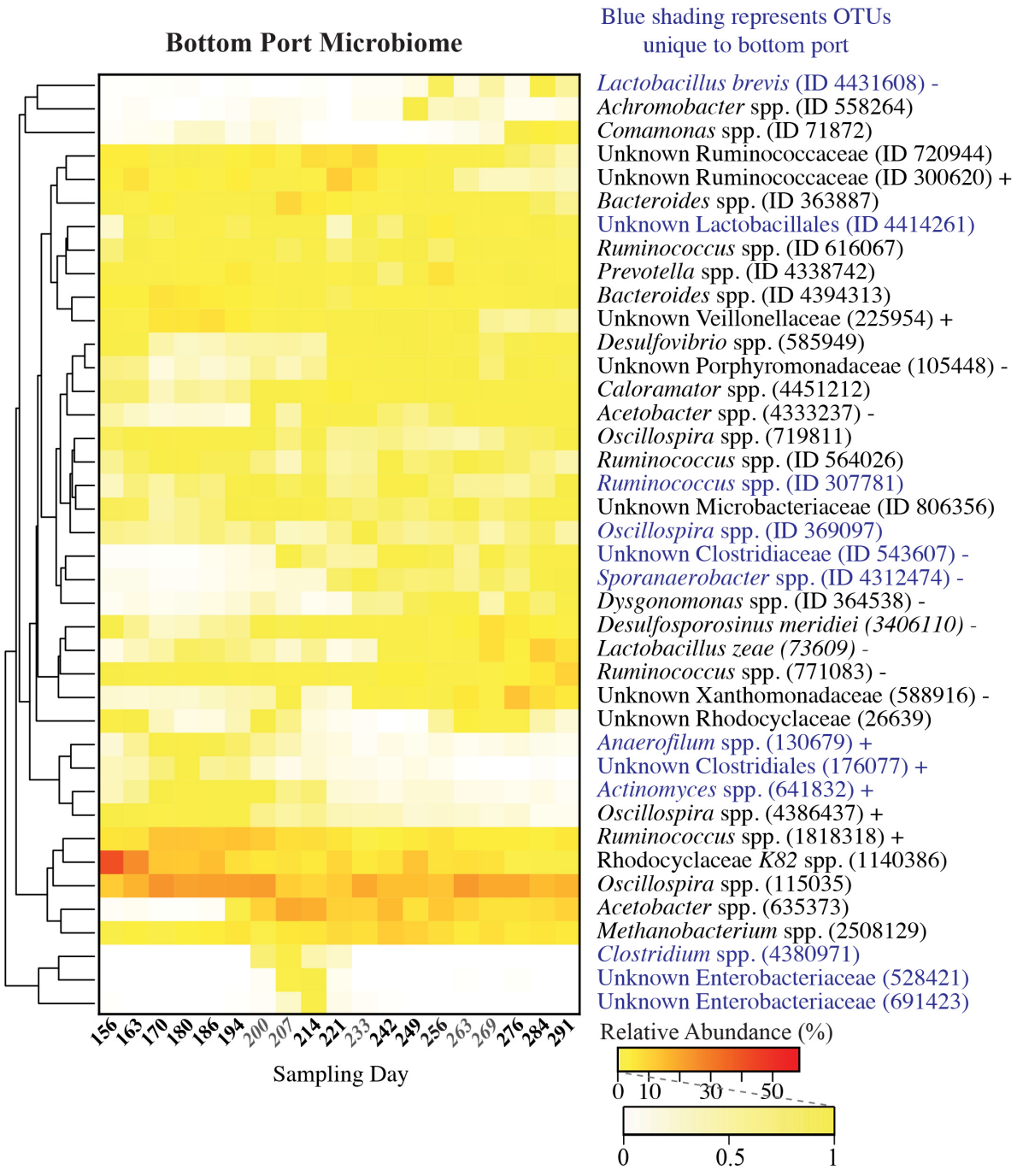

Figure 4.4. Heat map of relative OTU abundances in biomass samples taken from the bottom port of the bioreactor during the main periods of the study (Periods 1 to 7). Sampling day numbers are shaded grey to represent even periods in the study (Periods 2, 4, and 6; see Table 4.1). Relative abundance (%) is represented by the color gradient shown. OTUs that reached higher than 1% relative abundance in any one sample are represented, resulting in 40 OTUs. OTUs are clustered hierarchically (average linkage) based on the Bray-Curtis dissimilarity index. Lowest level taxonomy names as well as OTU IDs are provided. Blue shading represents OTUs that are unique

to bottom of the bioreactor (*i.e.,* not found in the samples taken from the middle of the bioreactor). + or – symbols represent whether the relative abundance of the OTU was found to be significantly positively (+) or negatively (-) correlated with *n*-caprylate specificities based on Spearman's rank correlation coefficient ($p<0.001$). *n*-Caprylate specificities were higher in the earlier periods (*i.e.,* sampling days 156 to 221 in the above figure) than in the later periods of the study.

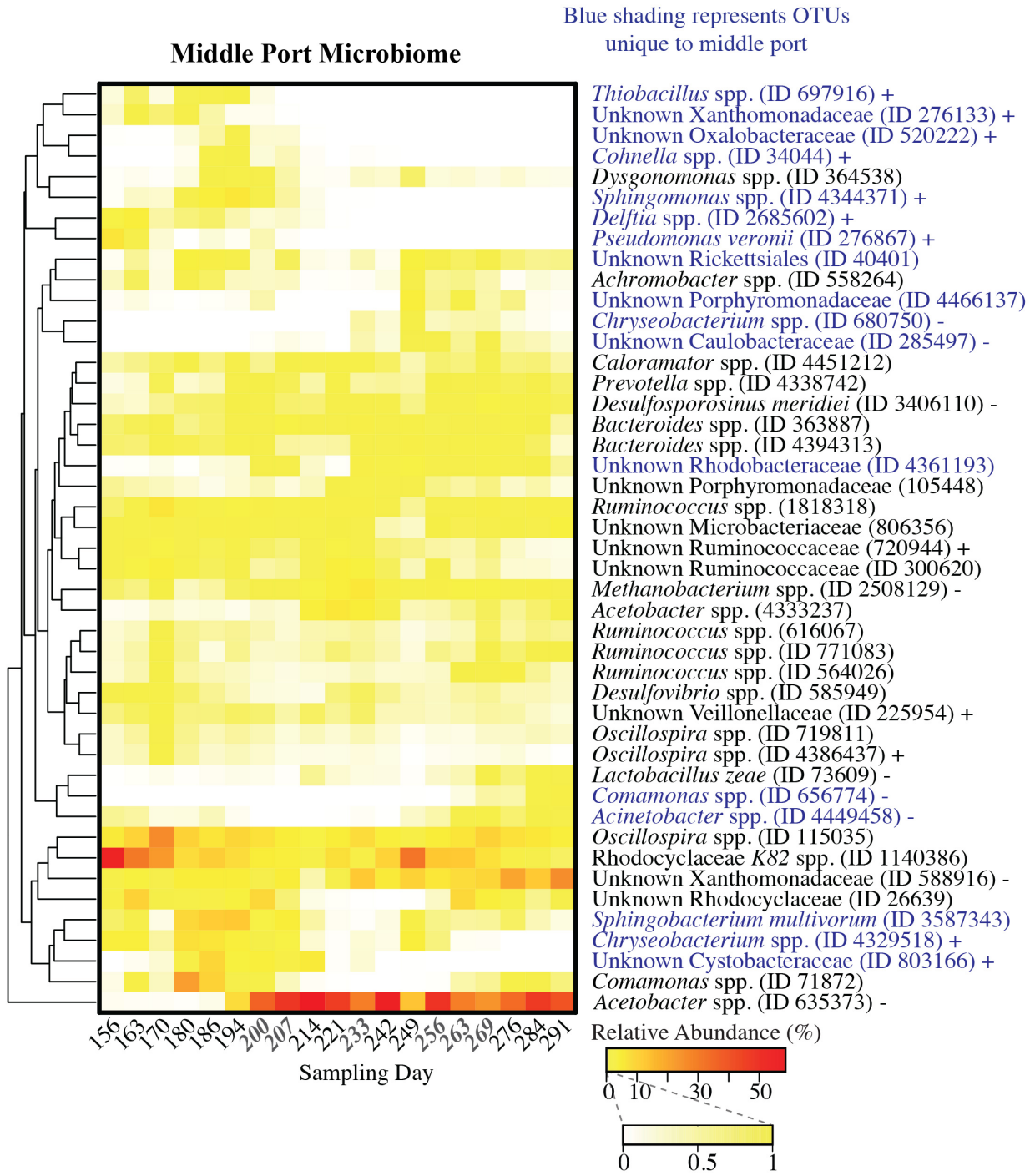

Figure 4.5. Heat map of relative OTU abundances in biomass samples taken from the middle of the bioreactor during the main periods of the study (Periods 1 to 7). Relative abundance (%) is represented by the color gradient shown. OTUs that reached higher than 1% relative abundance in any one sample are represented, resulting in 45 OTUs. OTUs are clustered hierarchically (average linkage) based on the Bray-Curtis dissimilarity index. Lowest level taxonomy names as well as OTU IDs are provided. Blue shading represents OTUs that are unique to middle of the bioreactor

(*i.e.,* not found in the samples taken from the bottom of the bioreactor). + or – symbols represent whether the relative abundance of the OTU was found to be significantly positively (+) or negatively (-) correlated with *n*-caprylate specificities based on Spearman's rank correlation coefficient ($p<0.001$). *n*-Caprylate specificities were higher in the earlier periods (*i.e.,* sampling days 156 to 221 in the above figure) than in the later periods of the study.

For both the bottom and the middle of the bioreactor, we examined which OTUs were positively or negatively correlated with *n*-caprylate specificities based on Spearmans rank coefficient ($p<0.001$). Some common patterns emerged between the two sampling locations. Two different OTUs belonging to the family Ruminococcaceae were positively correlated to *n*-caprylate specificities in one of the two locations (OTU ID 300620 in the bottom (**Fig. 4.4**); OTU ID 720944 in the middle (**Fig. 4.5**)). These OTUs were also found to be positively correlated with *n*-caprylate productivity in the *prior* study in our lab using the same bioreactor by Kucek et al. (2016c). In addition, a Veillonellaceae family OTU (ID 225954) and an *Oscillospira* OTU (ID 4386437), which belongs to the Ruminococcaceae family, were positively correlated to *n*-caprylate specificities in both locations sampled on our bioreactor (**Fig. 4.4 and 4.5**). In the study by Kucek et al. (2016c), a different *Oscillospira* genus member (ID 115035), which also appeared in our bioreactors (**Fig. 4.4 and 4.5**), was found to be correlated to *n*-caprylate productivity. Another OTU belonging to the *Anaerofilum* genus (ID 130679), which is also in the Ruminococcaceae family, was also positively correlated to *n*-caprylate specificities in our bioreactor, though it was only seen in the bottom bioreactor samples (**Fig. 4.4**). Members of the Ruminococcaceae family, specifically the *Clostridium* cluster IV have been found to be associated with: 1) *n*-caproate production in a bioreactor fermenting yeast fermentation beer (Agler et al. 2012a); 2) *n*-caprylate production in the study by Kucek et al. (2016c) that employed the same bioreactor setup as this study; 3) *n*-caproate, *n*-caprylate, and trace amounts of *n*-decanoate production from biorefinery

thin stillage (see paper by Andersen et al. (2017) that links an uncultured *Clostridium* species, which was most closely related to *Clostridium* sp. BS1-1 based on 16S analysis, to medium-chain carboxylate production rates); 4) *n*-butyrate production in the human gut (see review papers on *Oscillospira* spp. by Gophna et al. (2017) and Lopetuso et al. (2013)); and 5) *n*-caproate production from lactate in Chinese strong liquor fermentation, (see papers by Zhu et al. (2015a, 2017) that report isolating a Ruminococcaceae CPB6 strain). Most of the *Oscillospira* spp. members remained uncultured (Gophna et al. 2017). To our knowledge, Cuiv et al. (2015) was the first paper to report isolating an *Oscillospira* spp. (in a rumen fluid-containing media).

Common patterns were also observed between the bottom and middle of the bioreactor in terms of which OTUs were negatively correlated with *n*-caprylate specificities. Different OTUs classified as *Acetobacter* spp. (OTU ID 4333237 in the bottom (**Fig. 4.4**) and OTU ID 635373 in the middle (**Fig 4.5**)), a *Desulfosporosinus meridiei* OTU (ID 3406110), a *Lactobacillus zeae* OTU (ID 73609), and an OTU in the family Xanthomonadaceae (ID 588916) (**Fig. 4.4 and 4.5**) were negatively correlated with *n*-caprylate specificities. In the middle of the bioreactor, a *Methanobacterium* OTU (ID 2508129) was found to be negatively correlated to *n*-caprylate specificities (**Fig. 4.5**). *Acetobacter* is an obligate aerobe that can convert ethanol to acetic acid. In the study by Andersen et al. (2017), the authors also observed *Acetobacter* in their anaerobic bioreactors and speculated that their survival could be due to trace amounts of oxygen entering in the non-anaerobic feed. It is possible that a similar situation was happening in our study, since we did not attempt to make the feed line completely anaerobic. In the study by Andersen et al. (2017), the authors observed that when populations of *Lactobacillus* spp. and *Acetobacter* spp. declined, the overall medium-chain carboxylate productivity of the bioreactor declined, which is the opposite of what we observed in this study. Differences in the substrates used between the

Andersen et al. (2017) study, which used a more complex substrate thin stillage, and our study, in which we used ethanol, acetate, and basal media, may have contributed to this difference.

4.4.4 Bioreactor System Unable to Recover Performance After Low Substrate Ratio

Following the lowest substrate ratio 1.8 g COD ethanol to g COD acetate in Period 7, where no *n*-caprylate production was observed (**Table A2.4**), the bioreactor never recovered to its *prior* level of performance. In Period 8, the substrate ratio was increased to 5.1 (*i.e.,* similar to the ratio employed in Period 5), and the bioreactor was operated for a period of 45 days. With this increase in substrate ratio, *n*-caprylate production was again detected in the bioreactor. However, the *n*-caprylate productivity did not recover to previous levels (**Table A2.4**). Medium-chain conversion efficiency for this period was only 19.6 ± 2.1% compared to 46.6 ± 5.9% in Period 5 (**Table A2.3**). Similarly, in Periods 9 through 11, the bioreactor never returned to the previous levels of medium-chain conversion efficiency observed in the main periods of the study (**Table A2.3**). In Period 8, gas production began to increase (**Table A2.7**). Across the main periods of the study (Periods 1 to 7) the average gas production was 0.38 ± 0.01 mL d^{-1}. In Period 8, the gas production increased to 1.22 ± 0.06 mL d^{-1}, and by Period 11 the gas production had increased to 2.74 ± 0.04 mL d^{-1} (**Table A2.7**). The gas composition per period (*i.e.,* carbon dioxide, hydrogen, and methane) is reported in Table **A2.8**. It is possible that at the low substrate ratio in Period 7 (*i.e.,* 1.8), methanogens were able to take over the system and shift the bioreactor away from medium-chain carboxylic acid production, though we were unable to confirm this based on our 16S rRNA gene sequencing results (data not shown). Due to the failure of the system to recover, we were unable to repeat the experiment to confirm that higher ethanol-to-acetate ratios lead to higher *n*-caprylate

specificities.

Previously, the study by Kucek et al.(2016c) overloaded this bioreactor by increasing the organic loading rate to 63.8 g COD L^{-1} d^{-1}. Following this major disturbance, it took approximately half a year to recover the bioreactor performance (data not shown). The microbiome in this bioreactor is shaped to efficiently produce medium-chain carboxylates by lowering the pH and using product extraction. In a study by Agler et al. (2012a) that used a similar setup to produce medium-chain carboxylates from corn beer, it was observed that higher *n*-caproate productivity correlated with a more uneven microbiome. Similarly, the community in our bioreactor was significantly less diverse ($p<<0.0001$) than the anaerobic digesters discussed in Chapter 3, as measured by the Shannon coefficient index, which accounts for richness and evenness of the microbiome structure, as well as the number of distinct OTUs (richness) (**Table A2.6**). It is important to note, though, that the bioreactor in this study was fed a less complex substrate (*i.e.,* ethanol and acetate) as compared to the cow manure fed to the anaerobic digesters. Nevertheless, shaping the bioreactor to produce medium-chain carboxylates at high rates and efficiencies may come at the price of decreased bioreactor stability. Functional redundancy is defined as ability of certain members of a microbiome to take over the functions previously performed by negatively-affected members of the community following a disturbance (Allison and Martiny 2008). The importance of functional redundancy in maintaining the stability of anaerobic digester microbiomes was discussed in Chapter 3. To improve the stability of these chain-elongation bioreactors, work should be done to identify what microbes are functionally redundant in these systems (especially in regards to carrying out chain elongation) and to understand what operating conditions contribute to their survival and growth.

4.5 Conclusions

In conclusion, this study demonstrated that higher ethanol-to-acetate substrate ratios lead to higher *n*-caprylate specificities in a bioreactor with continuous product extraction. It is important to note that only one bioreactor was used in this study, which limits the conclusions that can be drawn. At very low ethanol-to-acetate substrate ratio (~1.8 based on COD), no *n*-caprylate production was observed. To our knowledge, this is the first study to demonstrate *n*-caprylate production from primarily ethanol in the substrate (*i.e.,* ethanol-to-acetate ratio greater than 200), though, a significant fraction of excess ethanol was wasted at this ratio. This study also characterized the microbiome present at the different substrate ratios tested and examined which microbes were correlated with improved *n*-caprylate specificity in the bioreactor. To improve the stability in chain-elongation systems, this points toward the need to identify and characterize what microbes are functionally redundant in these systems and what conditions optimize their growth.

4.6 Acknowledgements

The authors would like to thank Hanno Richter for assistance with the thermodynamic calculations. This work was supported by the U.S. Army Research Laboratory and U.S. Army Research Office under contract/grant number W911NF-12-1-0555 and the NSF SusChem Program (Award #1336186) awarded to L.T.A. C.M.S. acknowledges funding from the United States Environmental Protection Agency STAR Fellowship Assistance Agreement number FP-91763801-0.

CHAPTER 5

VARIABILITY OF THE MICROBIOME INVOLVED IN *N*-CAPRYLATE PRODUCTION IN ANAEROBIC BIOREACTORS WITH IN-LINE PRODUCT EXTRACTION

Adapted from: Spirito, Werner, Kircher, Guzman, Angenent; In Preparation for *Environmental Science and Technology*

5.1 Abstract

Chain-elongation microbiomes can be used to convert ethanol- or lactic-rich organic wastes to medium-chain carboxylic acids. This technology is mostly still at the laboratory and startup scale, and the microbiome involved in this process has not been fully characterized. For this technology to reach an industrial scale, an increased understanding of what contributes to the stability and performance of these systems is needed. This study sought to study the composition and function of the chain-elongation microbiome, in particular in response to changes in hydrogen partial pressures in the system. Here, three 5-L bioreactors were fed with ethanol and acetate, operated at a pH of 5.5, a temperature of 30ºC, and with inline product extraction, to promote the formation of *n*-caproate and *n*-caprylate. Illumina 16S rRNA gene sequencing, shotgun metagenomics, and metaproteomics were used to characterize the bioreactor microbiomes during the operating period. In the bioreactors, despite similar imposed operating conditions, we observed different microbial communities and different *n*-caprylate productivities. In the bioreactor with the lowest *n*-caprylate conversion efficiency, relatively high hydrogen partial pressures were observed. Subsequent sparging of nitrogen gas reduced the hydrogen partial pressure in that bioreactor and resulted in a

higher *n*-caprylate conversion efficiency. However, across all the bioreactors a common relationship between hydrogen partial pressures and *n*-caprylate conversion efficiency was not found. An *Oscillospira* OTU in the family Ruminococcaceae dominated in the bioreactor with the highest *n*-caprylate specificities and productivities. Metagenomics analysis revealed genes for enzymes involved in reverse beta-oxidation, though the full set of genes for reverse beta-oxidation was not found in any of the partial genomes recovered. This may indicate that the reference database used was missing the genes for our acclimated microbiomes. In conclusion, this study found that the chain-elongation microbiome was variable and unique.

5.2 Introduction

Laboratory studies have demonstrated efficient production of medium-chain carboxylates, including *n*-caproate, *n*-heptanoate, and *n*-caprylate, by anaerobic fermenter microbiomes at rates comparable to methane production by anaerobic digester microbiomes (Kucek et al. 2016c) (Grootscholten et al. 2013b). Medium-chain carboxylates can be produced *via* reverse beta-oxidation in which ethanol, lactic acid, or another electron donor is oxidized, providing energy and reducing equivalents for the chain elongation of short-chain carboxylates, such as acetate and *n*-butyrate, to longer chain carboxylates, such as *n*-caproate (6-carbon chain) and *n*-caprylate (8-carbon chain) (Cavalcante et al. 2017, Spirito et al. 2014). The bacterium that has been the most well-characterized in terms of carrying out reverse beta-oxidation is *Clostridium kluyveri*, with studies dating back to the mid 1900s of *C. kluyveri* carrying out chain elongation from ethanol and acetate (Barker et al. 1945, Bornstein and Barker 1948). The genome of *C. kluyveri* was sequenced in 2008 (Seedorf et al. 2008). Other bacteria have also been found to be capable of forming *n*-caproate or *n*-caprylate from a variety of substrates, including, but not limited to, *Megasphaera*

elsdenni (Jeon et al. 2016), *Peptococcus niger* (Ding et al. 2010), members of the genus *Eubacterium* (Wallace et al. 2004, Wallace et al. 2003), and *Clostridium* cluster IV members within the family Ruminococcaceae (Zhu et al. 2017). In recent studies in our lab with high rates of *n*-caproate and *n*-caprylate formation (Kucek et al. 2016b, Kucek et al. 2016c), we did not observe *Megasphaera elsdenii* or *Clostridium kluyveri*, though we did observe Ruminococcaceae family members, based on 16S rRNA gene sequencing of the bioreactor microbiome.

Hydrogen partial pressures can affect the product spectrum formed during primary fermentation (Andel et al. 1985, Bastidas-Oyanedel et al. 2012, De Kok et al. 2013, Zhang et al. 2013b) and chain elongation in anaerobic bioreactors (Arslan et al. 2012, Steinbusch et al. 2011, Yu and Mu 2006). Higher hydrogen partial pressures can serve to promote the formation of more reduced products, like *n*-caproate and *n*-caprylate, and to prevent oxidation of the produced carboxylates (Arslan et al. 2016, Weimer and Kohn 2016). Few studies have thoroughly investigated the range of hydrogen partial pressures that promote selective medium-chain carboxylate production in these systems. Papers by Arslan et al. (2012, 2013) investigated the effect of elevated hydrogen partial pressures (up to 0.2 MPa) on short and medium-chain carboxylate production. The authors found that *n*-caproate production was feasible even under 0.2 MPa hydrogen partial pressure (Arslan et al. 2013). They found that while pressurizing the bioreactor headspace with hydrogen increased total carboxylate production, but it did not lead to selective production of *n*-caproate (Arslan et al. 2012).

Here, we employed Illumina 16S rRNA gene sequencing, shotgun metagenomics, and metaproteomics to characterize the microbiomes in three bioreactors with in-line product extraction, which were producing *n*-caproate and *n*-caprylate. Despite similar imposed operating conditions in the first period of the study, we observed marked difference in the microbiomes of

the three bioreactors and in their medium-chain carboxylate productivities and selectivities. In addition, we observed some effects of different hydrogen partial pressures on the bioreactor microbiome composition and function, though our results were not conclusive.

5.3 Materials and Methods

5.3.1 Continuously Fed Bioreactor System

Three stainless-steel bioreactors of 5.5-L total volume (5-L working volume and 0.5-L headspace volume) were constructed. The internal diameter of the bioreactors was 10.2 cm (with the exception of the cone at the bottom of the bioreactors). We refer to these bioreactors as Reactor 1, Reactor 2, and Reactor 3. A photo and a diagram of the setup of the bioreactors can be found in **Fig. 5.1** and **Fig. 5.2,** respectively. The bioreactors were continuously stirred by recirculating the bioreactor broth at a rate of ~40 mL min^{-1} with a peristaltic pump (Cole Parmer, Part No. 7520-10) that removed bioreactor broth from near the top of the bioreactor liquid level and returned it to the bioreactor at the bioreactor base (internal recycle line). The temperature in the bioreactors was maintained at 30 ± 1ºC *via* a heating jacket. A pH probe (Omega, Part No. PHE-7353-15) was mounted in approximately the middle of the bioreactor. A controller (Cole Parmer, Part No. YO-56705-00) was used to maintain the pH in each of the bioreactors at ~5.5 *via* periodic additions of 0.5 M HCl with an acid-addition pump (Omega, Part No. PHP-601). The HCl was added at the base of the bioreactor where the continuously recycled bioreactor broth was returned to the reactor. Fresh media containing ethanol and acetate was continuously fed from a refrigerated vessel (4ºC) into the base of the bioreactor using a peristaltic feed pump. Another peristaltic pump was used to remove effluent from the bioreactor (at one-third of the way up the bioreactor) at approximately the same flow rate. Liquid level switches (Omega LV-11) were used to maintain the liquid level

in the bioreactor (*i.e.*, if the broth level became too high, the switches would turn off the effluent pump temporarily and *vice versa*). Gas exit lines from the top of the bioreactor led to a condensation trap, bubbler, and then a gas flow meter (Calibrated Instruments Inc., Ritter MilliGas Counter Series MGC-1 V3.1, Hawthorne, NY). The condensation trap and bubbler (which was filled approximately halfway with water) were made in-house with stainless-steel parts. All the tubing used in the bioreactor setup was stainless steel, with the exception of the tubing used in the peristaltic pumps, which was Viton and norprene.

Figure 5.1. Photo of the setup of the three stainless-steel bioreactors. From this angle, principally the membrane contactors and stainless-steel lines of the pertraction system can be seen.

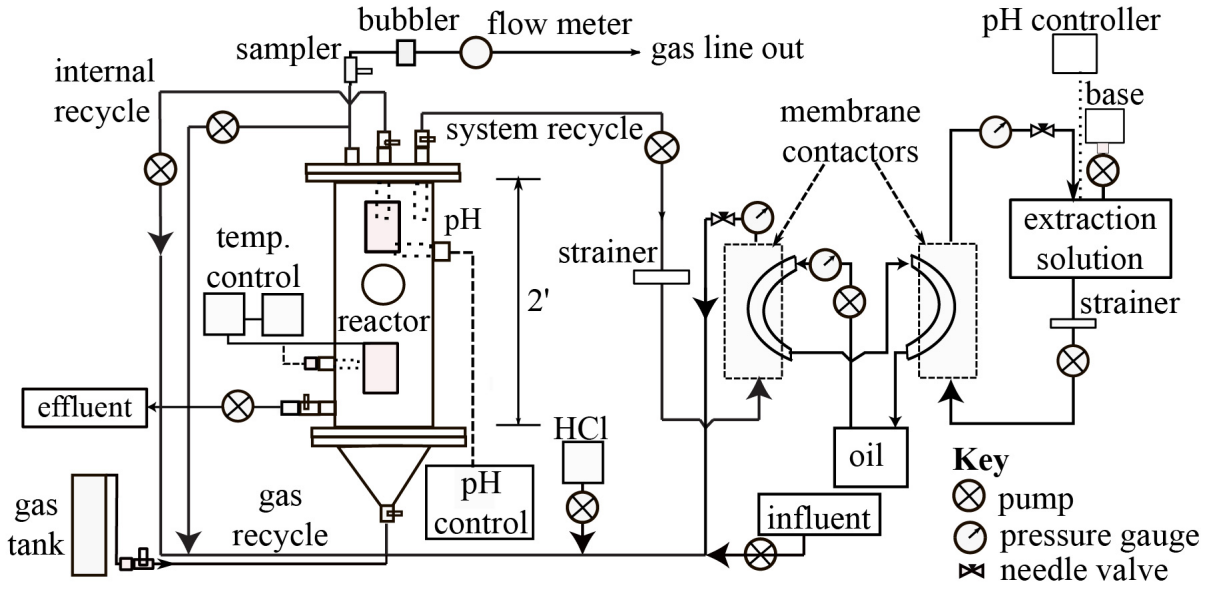

Figure 5.2. Diagram showing the setup of each bioreactor. Each bioreactor had temperature control and pH control. In addition, gas was sparged into the bottom of the bioreactor during Period 2. The pertraction system setup with the membrane contactors, oil, and alkaline extraction solution is shown to the right in the Figure.

5.3.2 Gas sparging setup

Stainless-steel lines were setup to sparge the bioreactor broth with either nitrogen, hydrogen, or both gases combined. The lines ran from tanks of nitrogen or hydrogen gas (pressure of tank set at 40 psi) to the bioreactor and entered the bioreactor near the bioreactor base. One-way check valves located near the bioreactor base were used to ensure uni-directional flow into the bioreactor. Gas flow meters (Cole-Parmer 65-mm Correlated Flowmeter, Part No. EX-03216-02), as well as needle valves, were used to control the flow rate of the gases into the bioreactor system. The gas flow rate entering the bioreactors was not measured. The gas flow rates leaving the bioreactors were measured and are reported in **Table 5.1**.

5.3.3 Pertraction system

For product extraction, we used a setup similar to the one previously described by Agler et al. (2012a). Forward and backward membrane contactors (8.1 m^2 each, Membrana Liqui-Cel 4 X 13, X50 Membrane, Charlotte, NC, USA) were connected to the bioreactor setup (**Fig. 5.1 and Fig. 5.2**). A peristaltic pump (Cole Parmer 7553-30) was used to recirculate bioreactor broth at a flow rate of 48 mL min^{-1} through the shell-side of the forward contactor. Broth was removed from near the top of the bioreactor, passed through a filter (McMaster Carr High-Pressure Stainless Steel Y-Strainer, 1/2 NPT Female, 100 Mesh) to remove particulate matter, and then passed through the contactor and back to the bioreactor. The filter was periodically cleaned to remove the built up particulate matter, on an approximately monthly basis. We should note that system recycle line that passed liquid through the contactor was different than the internal recycle line that was directly recirculating bioreactor broth. A mineral oil solvent with 30 g L^{-1} tri-*n*-octylphosphine oxide (TOPO) (Sigma Aldrich, St. Louis, MO, USA) was continuously circulated at a flow rate of ~30 mL min^{-1} (Cole Parmer 7553-30) through the lumen side of both the forward and backward membrane contactors. The purpose of the solvent was to primarily extract the more hydrophobic medium-chain carboxylates (as compared to the short-chain carboxylates) from the bioreactor broth. A well-mixed alkaline extraction solution was recycled at a flow rate of ~30 mL min^{-1} (Cole Parmer 7553-30) through the shell-side of the backward contactor. 0.3M sodium borate was initially used to buffer the extraction solution. The pH of the alkaline extraction solution was maintained at ~pH 9 *via* automated additions of 5 M NaOH using a pH controller and a corresponding base-addition pump.

5.3.4 Growth Media and Inoculum

The modified basal media was described previously (Kucek et al. 2016c, Vasudevan et al. 2014).

Ethanol and acetate were added to the media to achieve the desired organic loading rates. The pH of the media was adjusted with 5 M NaOH. The inoculum for the bioreactors was derived from a bioreactor microbiome that was fed with ethanol-rich yeast fermentation beer and had been operated as an anaerobic sequencing batch reactor for a period of approximately five years *prior* to the time the inoculum was collected (Agler et al. 2012a, Ge et al. 2015). The environmental conditions, such as temperature, pH, and product extraction, in the yeast fermentation beer bioreactor were similar to those used in this study. At the time of inoculation, 10% by volume (~500 mL) of the yeast fermentation beer bioreactor broth was added to the bioreactors in this study. The remaining volume consisted of basal media and substrate, ethanol, and acetate.

5.3.5 Experimental Periods for the Bioreactors

The main periods reported in this study are Periods 1 (days 75 to 141) and Period 2 (days 185 to 234). *Prior* to Period 1, there was a 75-day startup period for the bioreactors in which the organic loading rate was incrementally increased to the target loading rate of ~7 g COD L^{-1} d^{-1} at a hydraulic retention time of ~9 days. At the start of Period 1 (day 75), the biomass from all three bioreactors was combined together, mixed, and redistributed. During Period 1, the operating conditions (*i.e.*, temperature, pH, product extraction) were kept the same between the three bioreactors with the exception of small differences in the organic loading rates applied (**Table 5.1**). Between Periods 1 and Periods 2, gas sparging was tested out in the three bioreactors (*i.e.*, gas sparging was off and on between days 142 to 184). At the start of Period 2 (day 185), the biomass from the three bioreactors was again mixed and redistributed. In Period 2, Reactor 1 and Reactor 3 were continuously sparged with nitrogen gas, while Reactor 2 was continuously sparged with a mixture of hydrogen and nitrogen gas (**Table 5.1)**. During Period 1 when there was not sparging,

the gas flow rates measured leaving the bioreactors were very low (*i.e.,* no gas flow detected for Reactors 1 and 2 and an average flow rate of 0.502 ± 0.309 L d^{-1} for Reactor 3) (**Table 5.1**). Though the gas flow rate sparged into the bioreactors was not measured in Period 2, the sparging rates are assumed to be approximately equal to the measured exit gas flow rates reported in **Table 5.1** due to the low gas production rates observed in Period 1 without sparging. The bioreactors maintained pressure before the start of the operating periods. We assume that if leaks occurred they would be small.

Table 5.1. Operating data for the three bioreactors. HRT (hydraulic retention time), OLR (organic loading rate), and gas flow rate per bioreactor per period are reported. Gas flow rate was measured at the outlet of the bioreactor systems. The gas with which the systems were sparged is also indicated. Periods 1 and 2 are abbreviated as P1 and P2. Mean ± s.e. is reported.

	Period	Days	HRT (d)	OLR (g COD L^{-1} d^{-1})	Gas Flow Rate (L d^{-1})	Sparged with
Reactor 1	P1	75 to 141	9.1±0.3	7.3±0.5	0	No gas
	P2	185 to 234	9.0±0.3	7.6±0.3	24.702±13.732	N_2
Reactor 2	P1	75 to 141	9.2±0.5	7.5±0.5	0	No gas
	P2	185 to 234	8.9±0.3	7.3±0.4	6.137±7.001	N_2, H_2
Reactor 3	P1	75 to 141	7.9±0.3	7.9±0.3	0.502±0.309	No gas
	P2	185 to 234	8.4±0.3	7.5±0.3	10.186±4.827	N_2

5.3.6 Liquid and Gas Analysis

Liquid samples were collected from the bioreactor broth and the alkaline extraction solution every other day for the purpose of measuring carboxylate and ethanol concentrations. The 2-mL samples of the bioreactor broth were collected from a port in the system recycle line of the bioreactor, whereas the alkaline extraction solution samples were collected from a ~3-L well-mixed glass reservoir from which the extraction solution was being re-circulated. Samples were stored frozen

at -20ºC *prior* to analysis. Gas chromatography systems were used to determine the carboxylate and ethanol concentrations, as has been described by Usack et al. (2015).

Gas samples were collected from the gas exit lines of the bioreactors. Carbon dioxide, methane, and hydrogen concentrations (>0.2% by volume) were measured using a gas chromatography system, which has been described previously (Usack and Angenent 2015). A reduction gas detector (RGD) was used to measure hydrogen gas concentrations <0.2%, as has been described by Kucek et al. (2016c).

5.3.7 Total Protein Measurements

Total protein measurements were made of selected 2-mL liquid bioreactor broth samples. The total protein measurements were made using the Pierce BCA Protein Assay Kit (ThermoFisher Scientific, Catalog No. 23225). *Prior* to the measurements, the samples were thawed and centrifuged at 12,396 x *g* for 30 min. After discarding the supernatant, the samples were washed with phosphate buffer. Following this, NaOH was added to the samples, as well as to the BSA standard, to achieve a concentration of 1M NaOH. The samples and BSA standard were then boiled at 110ºC for 10 min on a heating block. After cooling on ice, dilutions of the samples and standard were prepared using 1M NaOH. Following the instructions of the Pierce BCA Protein Assay Kit, 25 µL of the diluted sample or standard were added to each well of the microplate (96 well flat bottom plate, Nunc) along with 200 µL of the working reagent prepared from the kit. Using a plate reader with Gen5 software, the microplate was shaken for 30 s and then incubated at 37ºC for 30 min in an external incubator. Following the incubation, the microplate was allowed to cool to room temperature and endpoint absorbance readings at a wavelength of 562 nm were made using the plate reader. A standard curve was constructed and used to calculate the total protein

concentrations (μg mL^{-1}) in the samples.

5.3.8 Calculations and Statistical Analysis of Operating Data

Carboxylate production rates are calculated as the average values for each operating period. Average effluent production rates per liter of bioreactor (g COD L^{-1} d^{-1}) and average transfer rates *via* product extraction (g COD L^{-1} d^{-1}) were summed to yield the total production rates per liter of bioreactor (g COD L^{-1} d^{-1}). COD stands for chemical oxygen demand. The average effluent production rates were calculated by dividing the average carboxylate concentration per period by the average HRT for that period. The average HRT per period was calculated based on the average effluent flow rate per period, which was determined gravimetrically. The average transfer rates were calculated by first plotting the increasing concentrations of the individual carboxylates in the alkaline extraction solution *vs.* time. Then the linear model function (lm) in the stats package in R v.3.3.2 (R Core Team 2016) was used to determine the slope and standard error of the best fit line through these points. The slope was then divided by the bioreactor working volume (5-L) to get the average transfer rate per period. Conversion efficiency is calculated as the individual carboxylate total production rate divided by the organic loading rate per period. Specificity is calculated as the individual carboxylate total production rate divided by the combined total production rate for all carboxylates during each period (where the carboxylates included are *n*-butyrate, *n*-caproate, and *n*-caprylate). Pertraction efficiency is calculated as the average transfer rate divided by the total production rate for each carboxylate. RStudio v.1.0.136 (RStudio Team 2016) was used to run all data analysis in R. All concentrations, rates, ratios, and efficiencies are reported as mean value ± standard error in the paper, unless noted otherwise in the text.

5.3.9 16S rRNA Gene Sequencing Analysis

Biomass samples for Illumina 16S rRNA gene sequencing analysis were taken from the internal recycle line of the bioreactors approximately weekly throughout the bioreactor operation. Approximately 10-mL of bioreactor broth was collected with a 60-mL plastic syringe and distributed into 2-mL Eppendorf tubes. These tubes were then centrifuged at 16,873 x *g* for four min and the supernatant was discarded. Pelleted biomass samples were stored at -80°C until further processing.

Genomic DNA was extracted using the PowerSoil-htp 96 Well Soil DNA Isolation kit (MO BIO Laboratories Inc., Carlsbad, CA) according to the protocol of the manufacturer. The DNA amplification protocol was described previously (Regueiro et al. 2015) with the following exceptions: Mag-Bind RxnPure Plus magnetic beads solution (Omega Biotek, Norcross, GA, USA) was used instead of Mag-Bind E-Z Pure and 50 ng DNA per sample was pooled instead of 100 ng. Duplicate PCR reactions of each DNA extract were performed and pooled *prior* to sequencing. Paired-end reads were joined in QIIME version 1.9.1 (Caporaso et al. 2010) using the joined_paired_ends.py script and then the joined reads were uploaded to QIITA (qiita.microbio.me) for further processing. The sortmerna method (Kopylova et al. 2012) was used to bin sequences in operational taxonomic units (OTUs) at 97% identity. Taxonomy was assigned for representative sequences selected for each OTU using the Greengenes v13.8 database from August 2013 (McDonald et al. 2012). The remaining analyses were performed locally in QIIME v1.9.1 and RStudio v1.0.136. Singletons were removed from the dataset resulting in 651 unique OTUs. Alpha diversity was analyzed the Shannon diversity index, Gini coefficient (measure of unevenness), and observed OTUs (richness) metric. 100 rarefactions at a depth of 6510 sequences per sample were performed and collated. The weighted Unifrac distance metric (Lozupone et al.

2006) was used to analyze beta diversity at an even sampling depth of 6510 sequences per sample. Statistical analyses of the alpha diversity results were performed using analysis of variance (ANOVA) and Tukey HSD in R. Principal coordinates analysis (PCoA) was used to visualize the differences in community between the samples. Heat maps were created to represent OTU relative abundance *via* the gplots package in R (Warnes et al. 2015).

Sequences were submitted to EBI under the accession number ERP024135. Sequences and study metadata are publically available in QIITA, which is an open-source microbiome storage and analysis resource, under study number 11227.

5.3.10 Shotgun Metagenomics Analysis

Biomass samples for shotgun metagenomics analysis were also collected from the internal recycle lines of the bioreactors. Samples were centrifuged, the supernatant was discarded, and the biomass was stored at -80°C until further processing. Genomic DNA was extracted using the PowerSoil DNA Isolation kit (MO BIO Laboratories Inc., Carlsbad, CA). A modified protocol was used as has been described by Kucek et al. (2016c). Picogreen (Quant-It Picogreen DSDNA, P7589, Fisher Scientific, Grand Island, NY) was used to measure the DNA concentration in the sample. Following this measurement, nine extracted DNA samples (three per bioreactor) were selected for sequencing. For Reactors 1 and 2, samples from days 137 (in Period 1) and days 199 and 218 (in Period 2) were sequenced. For Reactor 3, one of the time points sequenced was a combination of multiple sample days (days 137, 151, 154, 162) due to low DNA recovery from each of these time points. For Reactor 3, days 199 and 218 were also analyzed. DNA was stored at -80ºC *prior* to sequencing.

The nine metagenomic samples were barcoded and sequenced on two lanes (100 bp *per*

read; single-direction reads) using the Illumina HiSeq platform at the JP Sulzberger Genome Center at Columbia University (New York, New York), resulting in 477,902,544 reads. Raw sequencing reads have been uploaded to MG-RAST (Meyer et al. 2008) (study data available at: http://metagenomics.anl.gov/linkin.cgi?project=mgp21082). Reads were initially quality filtered to trim low-quality (qual "char" B) portions and remove trimmed reads <75 bp. This resulted in an overall raw dataset of 253,365,574 reads. The number of reads per sample, as well as the proportion that passed the quality filtering step can be found in **Fig. A3.1A**. The base pair frequencies at each base position through the raw reads, after the quality filtering, can be found in **Fig. A3.1B**. The full set of quality-filtered reads from all samples was then assembled *via* MegaHit v1.1.1-2-g02102e1 (Li et al. 2015a) using kmer lengths of 27 to 99 bp (step size 10 bp) and a minimum contig length of 500 bp, resulting in 225,200,668 total bp in the final assembly of 93,163 contigs, 223,750 bp maximum length, 2,417 mean length, and N50=5819 bp. Gene-coding regions were identified using MetaGeneMark (Zhu et al. 2010) and the resultant 271,182 protein sequences were searched against NCBI-nr using DIAMOND 0.8.36 blastp in sensitive mode (Buchfink et al. 2015), reporting the top 12% of hits in tabbed blast format. Consensus annotations were recorded for each protein in terms of function (SEED Subsystem role mapped from RefSeq IDs) and lowest common ancestor (LCA) taxonomy (from NCBI taxonomy) considering hits within 10% of the top bitscore, using custom perl scripts based on the method used in MEGAN (Huson et al. 2011). 247,896 of the 271,182 hypothetical protein sequences matched NCBI-nr reference sequences of known taxonomy, but only 10,008 of those proteins were assigned a SEED Subsystems role (function).

The relative abundance of each contig in each sample was estimated as normalized coverage. Megablast (Johnson et al. 2008) was used to map the raw reads to their respective

contigs, if any, (requiring 98% identity when aligning at least 75% of a query read length to the assembled contig). The query raw reads in this step were labeled with the sample name in the FASTA ID, following the QIIME (Caporaso et al. 2010) formatting style. Coverage in each sample was calculated based on the contig length and the overlapping alignment length from each of the Megablast results, and then these per-sample contig coverage values were normalized to 10 million reads per sample. The final data summary for comparative analysis was a table of annotated genes from the assembly, with attributes including the contig of origin, gene length, contig length, LCA taxonomy, SEED Subsystems role, EC number (if applicable), and normalized coverage in each sample; and, a table of contigs, their coverage values in each sample, and LCA taxonomy inherited from the gene data table (based on the highest frequency taxonomy string).

Assembled contigs were clustered based on their relative abundance profiles (coverage values, normalized to 10 million reads per sample) from all 9 time points using a Euclidian distance and an average neighbor clustering algorithm (Schloss et al. 2009) with a cutoff of 0.75 x coverage. Euclidian distance scales up with the number of samples (dimensions); in this case with 9 samples, a distance of 0.75 x covg translates to an average of about 0.25 x coverage per time point. This resulted in 2755 contigs (mean 4 contigs per cluster, max 539 contigs per cluster). Downstream analyses were performed on the 47 largest contig clusters, defined as clusters with over 500 genes, and the highest coverage clusters (114, mostly single contigs, had a total coverage across the nine samples of over 400 x (normalized to 10 million reads per sample). In the clusters with the highest coverage, 92 clusters (with primarily Bacteroidales members) had similar patterns across all the time points. Therefore, it was decided to manually cluster them by averaging their coverage per sample and renaming them as Bacteroidales_manual_cluster. This reduced the dataset to a total of 23 clusters.

5.3.11 Metaproteomics Analysis

Metaproteomics samples were collected from the internal recycle lines of the bioreactors. Approximately 200 mL of reactor broth was collected and distributed into four 50 mL centrifuge tubes. After centrifugation for 10 min at 8000 x *g* (at 4°C), the supernatant was discarded. The pellets were resuspended in a tris buffer solution and then redistributed to 2-mL Eppendorf tubes. These tubes were then spun at for 4 min at 16873 x *g*, and then the supernatant was discarded. The pelleted samples were then stored at -20°C until further processing.

Protein samples were extracted from bioreactor cell pellets (~100 μL bulk volume) using a gel-free, precipitation-free method to avoid loss of hydrophobic proteins. Cell pellets were suspended in 500 μL 50 mM Tris buffer (pH 8.0), and flash frozen 3 x with liquid nitrogen as an initial lysis step. Subsequently, 0.1% SDS, 10 mM NaCl, 0.02 M TCEP, and 2 M urea and were added and the sample was lysed by ultrasonication, on ice, at 60% amplitude for 5 min total pulse time, vortexed, and centrifuged 10 min at 12,000 x *g*. Half of the supernatant (~250 μL) was removed and saved. Then, to attempt to desorb more hydrophobic proteins from the pellet, 250 μL of acetonitrile was added to the cell pellet and remaining supernatant, vortexed, pelleted, and the supernatant from this step was then removed and re-combined with the first 250 μL of supernatant (the insoluble pellet was discarded), followed by reduction to approximately 400 μL total volume *via* speed vac. Total protein estimates were measured by Bradford assay. Protein samples were reduced with an additional 0.05 M TCEP in 0.1 M ammonium bicarbonate at 35°C for 1 h, alkylated with 40 mM iodoacetamide at room temperature for 30 min, and digested with Pierce Trypsin Protease MS-Grade at an estimated 1:20 trypsin:protein mass ratio for 12 h at 35°C with 1 mM $CaCl_2$. Sample protein precipitation was successfully avoided during digestion by diluting

the trypsin protease in 0.1 M ammonium bicarbonate buffer containing 0.02% SDS and 10% acetonitrile before combining with the protein sample. To quench digestion, samples were acidified to pH 3.5 with formic acid, acetonitrile was removed *via* speed-vac, acidified again to pH 3.5, and stored at -20 C. Tryptic peptides were purified using 1 mL Supelclean ENVI-18 SPE tubes and subsequently dissolved in 0.1% TFA / 0.5% acetonitrile for analysis by LC-MS.

LC-MS was performed using a Thermo Fisher UltiMate 3000 LC and LTQ-XL mass spectrometer with a standard ESI source. Microflow chromatography was performed on an Acclaim PepMap 100 column (1 mm x 15 cm; 3 um) at 40 μL/min using a 125 min gradient from 100 % water (1% formic acid) to 40 % acetonitrile. The LTQ-XL was operated in 3x double play mode with a 10 s dynamic exclusion time and CID activation.

5.4 Results and Discussion

5.4.1 Performance of Bioreactors Differed Despite Similar Imposed Operating Conditions

At the start of Period 1 (day 75), the broth from all the bioreactors was mixed and the bioreactors were restarted. During Period 1, all imposed operating conditions, such as organic loading rate, pH, and temperature, were kept the same or similar between the three bioreactors (**Table 5.1**). Despite this, the performance of the three bioreactors varied. Reactor 1 (R1) and Reactor 2 (R2) did have similar overall medium-chain conversion efficiencies, 58 ± 5% and 60 ± 5%, respectively (**Fig. 5.3A, Table A3.1**). However, Reactor 1 exhibited a much higher selectivity for *n*-caprylate production (**Table A3.2 and A3.3**), which resulted in a *n*-caprylate conversion efficiency of 53 ± 5% (**Fig. 5.3B**, **Table A3.1**). In contrast, in Reactor 2, the *n*-caprylate conversion efficiency was 38±4% (**Fig. 5.3B, Table A3.1**). As compared to Reactors 1 and 2, Reactor 3 performed poorly in Period 1 with a *n*-caprylate conversion efficiency of 18±1% and an overall medium-chain

conversion efficiency of 31±2% (**Fig. 5.3A and B, Table A3.1**). We should note that since the organic loading rates employed in the three bioreactors were approximately the same (**Table 5.1**), the conversion efficiencies reported are directly proportional to the carboxylate production rates in the bioreactors (for production rates see **Table A3.3**). In Period 1, differences were also noted in the carboxylate (*i.e.*, acetate, *n*-butyrate, *n*-caproate, and *n*-caprylate) concentrations in the effluent of the bioreactors (**Fig. 5.4A to C, Table A3.4**).

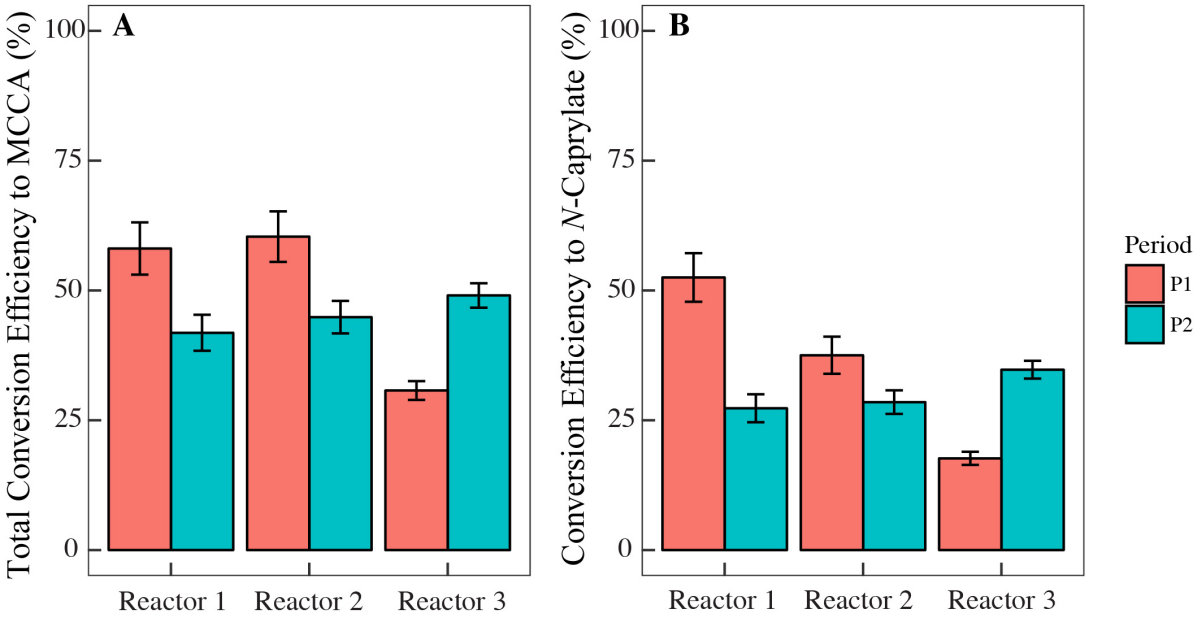

Figure 5.3. Total conversion efficiency of substrate to medium-chain carboxylates (*i.e.,* both *n*-caproate and *n*-caprylate) (**A**) and conversion efficiency to only *n*-caprylate (**B**) for the three bioreactors for Periods 1 and 2 (P1 and P2). Error bars represent standard error.

Though the imposed operating conditions were the same in Period 1, hydrogen partial pressures were higher in R3 and may have contributed to the worse performance. Across Period 1, the gas in the headspace of Reactor 3 contained 31.3 ± 9.6% hydrogen (by volume), whereas in Reactor 1 hydrogen was 11.1 ± 4.0% of the total gas and in Reactor 2 it was only 2.1 ± 1.8% (**Fig.**

A3.2A-C; **Table A3.5**). Hydrogen was produced in all bioreactors *via* the reverse beta-oxidation pathway. It is possible that Reactor 3 had more gas-tight connections than the other bioreactors, which limited hydrogen gas leakage from this bioreactor more than in the other bioreactors. A review paper by Angenent et al. (2004) notes that hydrogen can often be greater than 30% in the headspace of biohydrogen-producing reactors. A study by Zoetemeyer et al. (1982) observed up to 70% hydrogen in the headspace of an acidogenic bioreactor fed with glucose. In Period 2, nitrogen gas sparging into Reactor 3 lowered the percent hydrogen in the headspace of the bioreactor to 7.3 ± 4.6% (**Table A3.5**). During this period, the bioreactor performance improved resulting in an overall medium-chain carboxylate conversion efficiency of 49 ± 2% (**Fig. 5.3A**) and a *n*-caprylate conversion efficiency of 35 ± 2% (**Fig. 5.3B**).

Under the relatively high hydrogen partial pressures in Period 1 in Reactor 3, a significant fraction of the ethanol fed to Reactor 3 was leaving in the effluent, which resulted in an average effluent ethanol concentration of 163.9 ± 9.9 mM (**Fig. 5.4F**; **Table A3.4**). This ethanol concentration was higher than the concentrations leaving in the effluent of Reactors 1 and 2 in Period 1 (*i.e.,* 47.5 ± 4.1 mM and 29.1 ± 4.3 mM, respectively) (**Fig. 5.4D and E**, **Table A3.4**). As these were continuously-stirred systems, the concentrations measured in the effluent can be considered to be approximately equal to what the bioreactor microbiomes observed. It is likely that the high hydrogen partial pressures in Reactor 3 were limiting the rate of oxidation of ethanol to acetate within the reverse beta-oxidation pathway in this bioreactor. At a pH 5.5 and temperature, 30°C, we calculated that hydrogen partial pressures must be lower than approximately 1.013×10^{-3} MPa for uncoupled ethanol oxidation to occur in our system. However, ethanol oxidation could occur if it was coupled to chain elongation, though the rate of the reaction would likely be lowered (see Chapter 2). In Period 2, when the percent hydrogen in the headspace

of Reactor 3 was lowered to 7.3 ± 4.6% (**Table A3.5**) by sparging nitrogen gas, the concentration of ethanol leaving in the effluent of the bioreactor decreased to 98.2 ± 5.9 mM (**Fig. 5.4F**, **Table A3.4**) and the medium-chain carboxylate conversion efficiency improved due to the improved kinetics (**Fig. 5.3A**). The measured ethanol concentration in all the bioreactors were lower than concentrations that have been previously been found to be inhibitory in a pure (*i.e.,* 460 mM (Weimer and Stevenson 2012)) and in a mixed-culture study (*i.e.,* ~300 mM (Kucek et al. 2016c)). However, based on the significant fraction of ethanol leaving in our effluent, it appears that we may have been overloading our system.

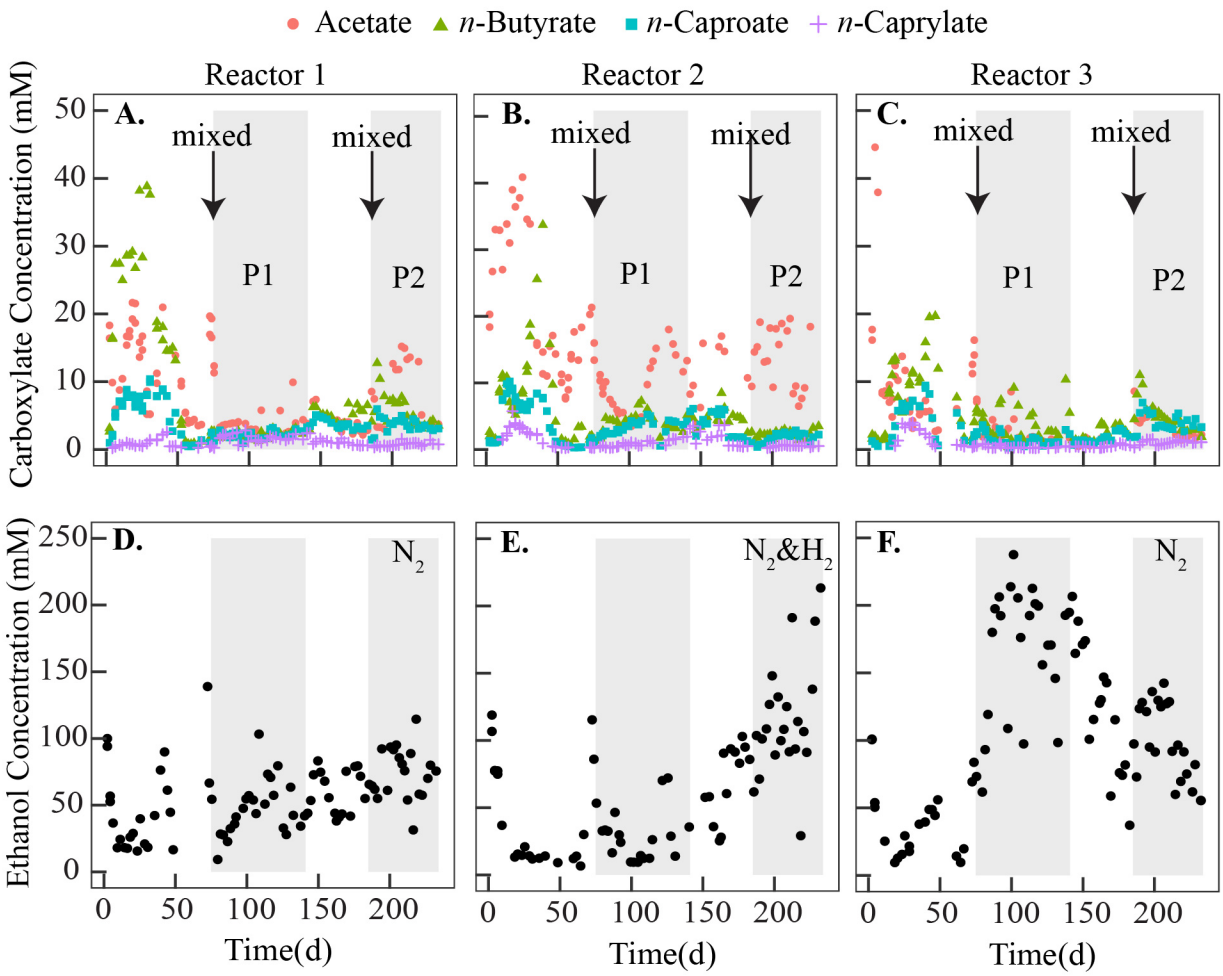

Figure 5.4. Effluent carboxylate concentrations over the operating period for Reactor 1 (**A**), Reactor 2 (**B**), and Reactor 3 (**C**). The two time points at the start of Periods 1 and 2 when the biomass from all the bioreactors was mixed together is indicated on these panels. Effluent ethanol concentrations over the operating period for Reactor 1 (**D**), Reactor 2 (**E**), and Reactor 3 (**F**) are also shown. The type of gases that were sparged into these bioreactors in Period 2 are indicated in these panels. Dark shading in panels **A-F** indicate Periods 1 and 2 (P1 and P2), which are the main periods reported on in this text.

5.4.2 Relationship between hydrogen partial pressure and medium-chain carboxylate production requires further study

Despite the apparent effect of hydrogen partial pressure on medium-chain carboxylate production

rate that we observed in Reactor 3, across all the bioreactors the relationship was not clear-cut. For Reactor 2, the medium-chain conversion efficiency was also lower under higher hydrogen partial pressures (*i.e.,* 45 ± 3% medium-chain conversion efficiency when hydrogen made up 13.2 ± 5.3% of the headspace gas *vs.* 60 ± 5% medium-chain conversion efficiency at an average hydrogen percent composition of 2.1 ± 1.8%; see **Fig. 5.3A, Table A3.1 and A3.5**). However, in Period 1, at a similar hydrogen partial pressure at which Reactor 2 experienced its worse performance, Reactor 1 experienced its best performance. For Reactor 1 in Period 1, hydrogen made up 11.1 ± 4.0% of the headspace gas. During this period, the MCCA conversion efficiency was higher than at the lower hydrogen percent composition of 0.7 ± 0.2% experienced by the Reactor 1 in Period 2 (**Fig. 5.3A, Table A3.1 and A3.5)**.

There are several possible reasons to explain the discrepancy in the results. In Period 2, Reactor 1 was sparged with nitrogen gas to lower the hydrogen partial pressure in the headspace. Sparging (*vs.* not sparging in Period 1) the bioreactor added another variable to the experiment. In Period 2, the gas flow rate out of Reactor 1 was the highest of all three bioreactors (*i.e.,* 24.702±13.732 L d^{-1}; **Table 5.1**). The gas sparging may have decreased biomass growth in the bioreactor. Total protein measurements of Reactor 1 show that the total protein concentrations in Reactor 1 were much lower in Period 2 than in Period 1 (*i.e.,* 796.84±104.46 (standard deviation shown, n=3) *vs.* 183.66±41.48 (standard deviation shown, n=4), respectively; **Fig. A3.3A and Table A3.6**). Another confounding factor in this study is that hydrogen partial pressures measured in the headspace of the bioreactor are only an approximation of the concentration of dissolved hydrogen actually experienced by the microbes in the bioreactor. Previous studies have found that dissolved hydrogen is often oversaturated in the liquid of the bioreactor and equilibrium with the gas headspace cannot be assumed (Beckers et al. 2015, Kraemer and Bagley 2006, Pauss et al.

1990). A dissolved hydrogen probe in the bioreactors would have allowed for more accurate measurement of the hydrogen concentrations experienced by our microbes. Finally, due to difficulty maintaining steady-state sparging rates with our bioreactor setup, there was wide variability in the hydrogen gas composition measurements for each bioreactor per period (**Table A3.5**). Thus, for the above reasons, it is difficult to make a definitive conclusion about the effect of hydrogen partial pressures on medium-chain carboxylate production. However, in general, higher hydrogen partial pressures appear to slow down the reverse beta-oxidation pathway, which results in a lower rate of coupled ethanol oxidation with a high concentration of ethanol leaving in the effluent and a lower medium-chain conversion efficiency as a result.

5.4.3 Microbial community differences between bioreactors

The bioreactors were all fed with the same substrate, ethanol and acetate, operated at similar loading rates (**Table 5.1**), and all produced *n*-caprylate (**Table A3.3**). In addition, the biomass from all bioreactors was mixed and redistributed at the start of Period 1. Despite this, the microbiomes of the three bioreactors were different. Principal coordinates analysis of the bioreactor microbiome samples shows a clear difference between the microbiomes, especially in regards to the communities of Reactor 1 *vs.* Reactor 2 in Period 1 (**Fig.** A**3.4A and B**). The PCoA also reveals that all the bioreactor microbiomes changed between Periods 1 and 2 (**Fig. A3.4A and B**). For Reactor 1, the dominant phylum in both Periods 1 and 2 was Firmicutes (**Fig. A3.5A**). For Reactor 2, Proteobacteria dominated in Period 1 and in Period 2, Proteobacteria shared approximately equal dominance with Bacteroidetes and Firmictures (**Fig. A3.5B**). For Reactor 3, Proteobacteria dominated in Period 1, while Bacteroidetes dominated in Period 2 (**Fig. A3.5C**). Heat maps of the OTUs that reached higher than 1 percent relative abundance in the bioreactors

show that while there were shared OTUs between the bioreactors, the dominant OTUs were different (**Fig. A3.6, A3.7, and A3.8**). A recent study by Andersen et al. (2017) on two identically-operated, chain-elongation bioreactors processing thin stillage observed that the two bioreactors developed different microbial communities, based on relative abundance data. The results of our study and the study by Andersen et al. (2017) are in contrast to studies in anaerobic digesters (Vanwonterghem et al. 2014a, Werner et al. 2014), which have found that replicate anaerobic digesters generally maintain similar microbial communities and functions over time. Overall, as compared to the anaerobic digesters reported on in Chapter 3 and the chain elongation bioreactor in Chapter 4, the bioreactors in this study had a relatively low diversity, as measured by the Shannon diversity coefficient, and were dominated by a few OTUs (**Table A3.7**). In addition, it was observed that, for Reactors 1 and 2, the Shannon diversity was significantly lower ($p<<0.0001$) in Period 1 than in Period 2 (**Table A3.7**). For both bioreactors, Period 1 was the period of best performance (as measured by medium-chain conversion efficiency; **Fig. 5.3A and Table A3.1**).

Results of the shotgun metagenomics analysis backed up the Illumina 16S rRNA gene sequencing results. A heat map of the assembled clusters from the shotgun metagenomics samples, which had over 400 x total coverage across the nine samples (normalized to 10 million bp per sample), shows that clusters identified as either unknown bacteria, Clostridiales, or more specifically as *Oscillibacter* genus members, were more dominant in the Reactor 1 samples, especially day 137 in Period 1 (**Fig. 5.5**). Clusters identified as *Acetobacter* or Acetobacteraceae were more dominant in Reactor 2, while clusters identified as Bacteroidales were more dominant in Reactor 3. Note that one of the clusters that was more dominant in Reactor 3 is the manually created cluster discussed in the methods section, which principally contained Bacteroidales contigs. Further work could be done to confirm this cluster by comparing the similarity of the GC

content of these contigs. The metagenomics samples were collected from the bioreactors at the end of Period 1 and at two time points in Period 2. At similar time points in the Illumina 16S rRNA gene sequencing data, we saw that an OTU classified to the genus *Oscillospira* (ID 526665) dominated Reactor 1 (**Fig. 5.6A**), an OTU classified to the genus *Acetobacter* (ID 635373) dominated Reactor 2 (**Fig. 5.6B**), and an OTU (ID 139337) classified to the family Bacteroidales dominated Reactor 3 (especially in Period 2 for Reactor 3) (**Fig. 5.6C**). The *Acetobacter* OTU and the Bacteroidales OTU were likely not directly responsible for *n*-caprylate production in the bioreactors. Bacteriodales members are primary fermenters (Jaenicke et al. 2011, Narihiro et al. 2015), whereas *Acetobacter* spp. can produces acetate from ethanol (Andersen et al. 2017). Notably, the *Acetobacter* spp. OTU was also found in the experiment reported in Chapter 4 in an upflow anaerobic filter bioreactor fed with ethanol and acetate, but its presence was correlated with reduced *n*-caprylate specificity.

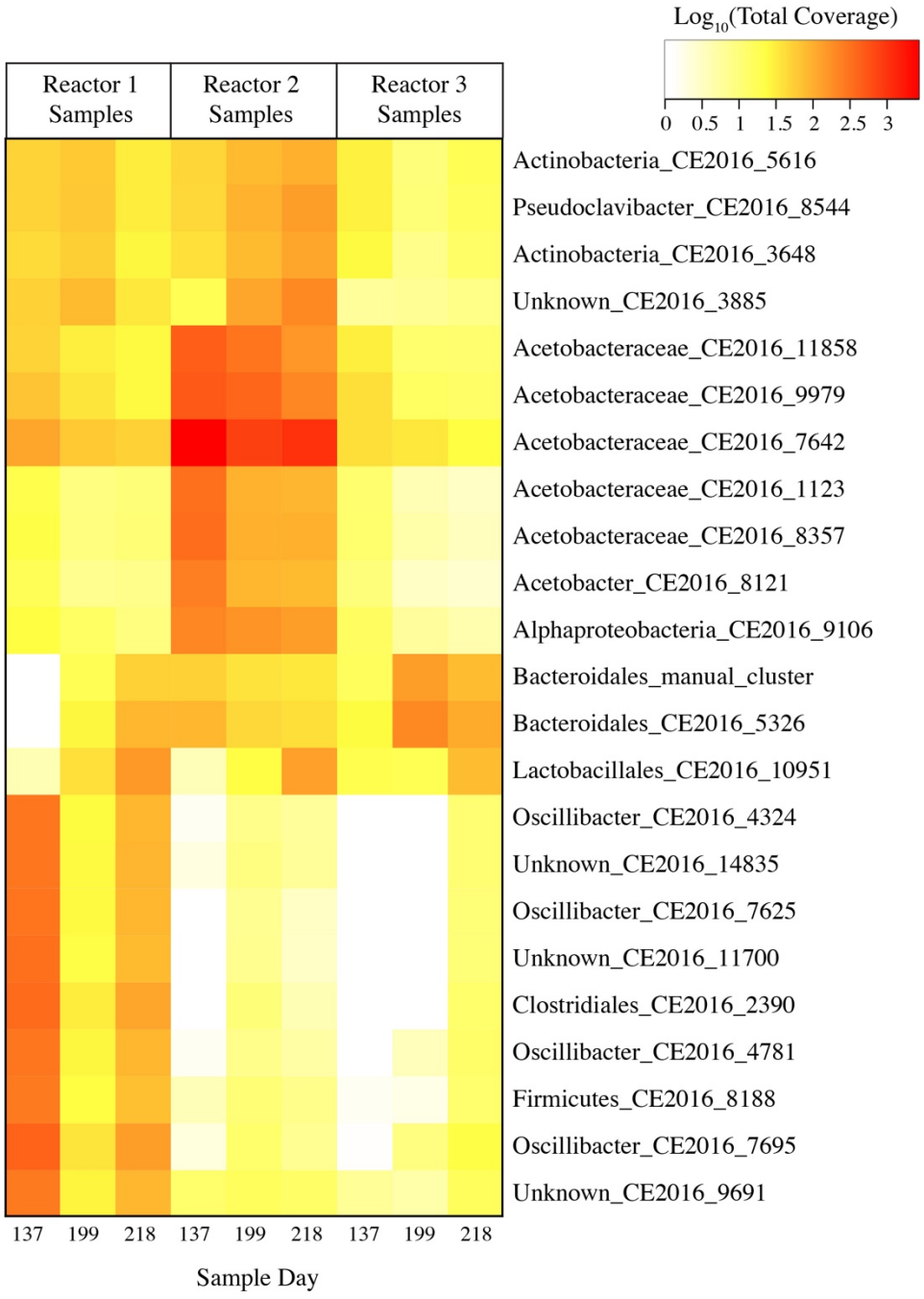

Figure 5.5. Highest coverage clusters (23 clusters, mostly single contigs) that had a total coverage across the nine shotgun metagenomics samples of over 400 x (normalized to 10 million reads per sample). The Bacteroidales manual cluster is the result of collapsing 92 clusters, which had similar abundance profiles across the samples into one manual cluster. The bioreactor from which the samples came from is indicated.

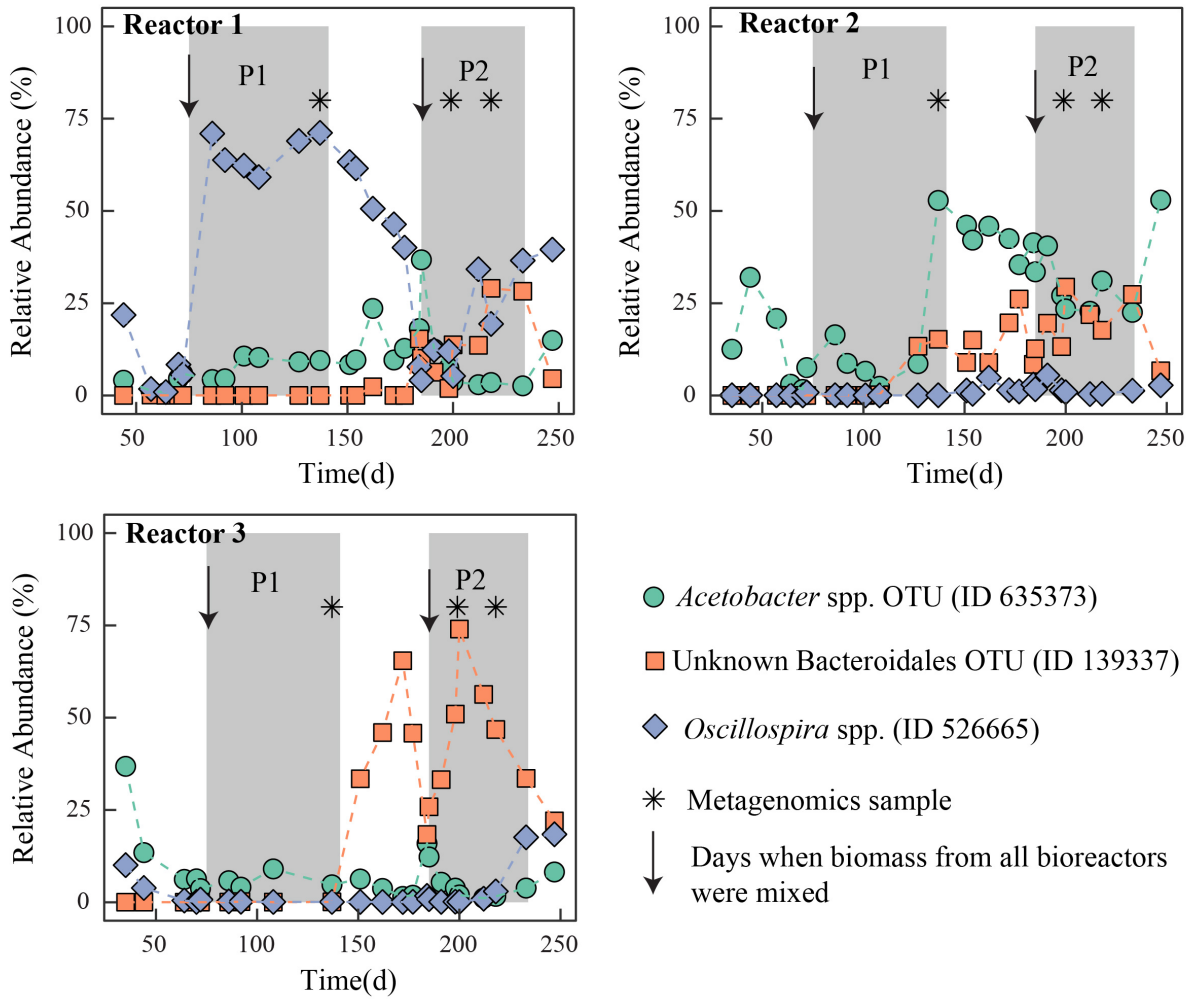

Figure 5.6. Relative abundance of dominant OTUs highlighted in the text in **A**) Reactor 1, **B**) Reactor 2, and **C**) Reactor 3 based on the Illumina 16S rRNA gene sequencing results. Periods 1 and 2 are shaded and are indicated on the figure as P1 and P2. The time points of the metagenomics samples are also indicated (*), as well as the dates when the bioreactors were mixed (arrow). Note that the first metagenomics sample for Reactor 3 was a combination of multiple sample days (days 137, 151, 154, 162) – only day 137 is indicated on the above Figure. Continuous gas sparging into the three bioreactors occurred in Period 2.

5.4.4 Oscillospira spp. *OTU dominates during period with high* n*-caprylate specificity*

The highest rate of *n*-caprylate production occurred in Period 1 in Reactor 1 (*i.e.,* 3.83 ± 0.23 g COD L^{-1} d^{-1}; see **Table A3.3**), where the *n*-caprylate conversion efficiency was 53 ± 5% and the overall medium-chain carboxylate conversion efficiency was 58 ± 5% (**Fig. 5.3A and B**).

Due to the high *n*-caprylate selectivity, the *n*-caprylate-to-*n*-caproate productivity ratio during this period was 9.4 ± 0.6 (g COD basis) (**Table A3.2**). During Period 1, an *Oscillospira* spp. OTU (ID 526665) dominated Reactor 1 (**Fig. 5.6A**). This *Oscillospira* OTU had 96% ID to an *Oscillibacter valericigenes* Sjm18-20 strain whose complete genome has previously been sequenced (Katano et al. 2012). As we discussed in Chapter 4, *Oscillospira* species are known to produce *n*-butyrate in the human gut (Gophna et al. 2017) and members of the family in which they fall (Ruminococcaceae) have been isolated from bioreactors producing *n*-caproate from lactate (Zhu et al. 2017, Zhu et al. 2015b). In addition, Illumina 16S rRNA gene sequencing studies have found Ruminococcaceae members associated with medium-chain carboxylate production in a bioreactor processing yeast fermentation beer (Agler et al. 2012a), a bioreactor fed with ethanol and acetate (Kucek et al. 2016c), and a bioreactor processing thin stillage (Andersen et al. 2017).

5.4.5 Low Percentage of Proteins Assigned Functions

As we mentioned in the methods section, in our metagenomics analysis, 247,896 of the 271,182 hypothetical protein sequences matched NCBI-nr reference sequences of known taxonomy, but only 10,008 of those proteins were assigned a SEED Subsystems role (function). The hypothetical protein sequences were good hits to the database, they had high percent identities, mostly full-length protein alignments, high bit scores and low e-values (**Fig. A3.9A-D**). However, a lack of functional annotations of the reference proteins that matched our hypothetical protein sequences in the NCBI-nr database limited our ability to assign functions to the protein-coding genes in the metagenome. It appears that our chain-elongation microbiome may be a unique system with many members that have yet to be isolated and characterized.

Despite this, we did find genes for enzymes involved in the reverse beta-oxidation pathway

in our metagenome (data not shown), though, we did not find all the genes for the full reverse beta-oxidation pathway in any of the partial genomes. Potentially, syntrophic organisms are involved in carrying out chain elongation in these bioreactors. However, the more likely scenario is that the reference database used was missing the genes of our acclimated microbiomes. We found two genes for butyryl-CoA dehydrogenase (EC 1.3.99.2), a gene for acetyl-CoA acetyltransferase (EC 2.3.1.9), and a gene for enoyl-CoA hydratase (EC 4.2.1.17). Based on the proteomics analysis, the butyryl-CoA dehydrogenase enzyme, which is responsible for conversion of crotonyl-CoA to butanoyl-CoA, was active in multiple time points across the three bioreactors (data not shown). The acetyl-CoA acetyltransferase enzyme, which is responsible for the conversion of acetyl-CoA to acetoacyl-CoA, was present in two time points from Reactor 2 and one from Reactor 3. Finally, the enoyl-CoA hydratase enzyme, which converts 3-hydroxybutanoyl-CoA to crotonyl-CoA, was active in two samples from Reactor 2 (data not shown).

5.5 Conclusions

In conclusion, this study characterized the microbiomes of three bioreactors fed with ethanol and acetate that were selectively producing medium-chain carboxylates. Illumina 16S rRNA gene sequencing, shotgun metagenomics, and metaproteomics methods were utilized. This study found the chain elongation microbiome to be both variable and unique. It was found that even under similar imposed operating conditions, the bioreactor microbiomes composition and performance were markedly different, which could be partially attributed to the buildup of hydrogen gas in one bioreactor. The most productive bioreactor had a relatively uneven microbiome dominated by an OTU in the genus *Oscillospira*. Shotgun metagenomics and metaproteomics analysis revealed genes for reverse beta-oxidation enzymes were present in the bioreactors and being expressed.

However, the complete reverse beta-oxidation pathway was not found. The genes responsible for reverse beta-oxidation in our bioreactor microbiomes appear to not be present in the reference database, which indicates that our acclimated chain-elongation bioreactor microbiome is unique.

5.6 Acknowledgements

The authors would like to acknowledge Chase Brett and Doug Caveney for all their help with the construction of the bioreactor. Thank you to Alex Marzelli for helping to prepare media for the bioreactor multiple times. Thank you to Jiajie Xu and Lauren Harroff for helping with bioreactor maintenance. We also acknowledge funding from US EPA STAR grant fellowship, as well as the U.S. Army Research Laboratory and U.S. Army Research Office under contract/grant number W911NF-12-1-0555.

CHAPTER 6

SUMMARY AND RECOMMENDATIONS FOR FUTURE WORK

6.1 Summary

The carboxylate platform utilizes anaerobic reactor microbiomes to convert organic wastes into energy and industrial chemicals. Anaerobic digestion is an established technology within the carboxylate platform in which a consortium of bacteria and archaea break down complex organic wastes to ultimately gaseous methane. Recently, a newer technology, called chain elongation, has been established within the carboxylate platform. In this technology, the anaerobic reactor microbiome is directed to produce medium-chain carboxylates, such as *n*-caproate and *n*-caprylate, instead of methane. This can be accomplished by lowering the reactor pH and carrying out continuous product extraction. In this work, I report on three time-series studies that examined the microbiome composition and performance of carboxylate platform bioreactor systems. The ultimate goal of these studies was to gain a better understanding of what contributes to stability in these systems and how the microbiomes in these systems change in response to perturbations and imposed parameters. In Chapter 3, I report on a study examining the effect of the dairy antibiotic monensin on anaerobic digester microbiomes. I found that monensin decreased the methane yield and led to some common shifts in the primary fermenter and syntrophic microbial populations in all anaerobic digesters in the study. However, in the anaerobic digesters that received manure from the monensin-dosed cows, the decrease in methane yield could be partially attributed to the differences in the compositions of the manure from the monensin-dosed cows *vs*. the control cows rather than a direct effect of the monensin antibiotic. I also observed that slow introduction of

relatively high concentrations of monensin into an anaerobic digester allowed for adaptation of the anaerobic digestion microbiome to the antibiotic by shifting its microbial community structure, whereas rapid introduction led to a rapid decrease in performance (*i.e.,* methane yield) of the anaerobic digester system. This study demonstrated the importance of functional redundancy in maintaining the stability of the anaerobic digester microbiome. In Chapter 4, I carried out a study looking at the effect of ethanol-to-acetate substrate ratios on the production rates of *n*-caprylate in a chain-elongation bioreactor (0.7-L working volume) with product extraction. Higher ethanol-to-acetate ratios led to higher *n*-caprylate specificities in the products of the bioreactor. In addition, *n*-caprylate production with mainly ethanol in the substrate was demonstrated. The microbiome was found to be relatively unstable and was sensitive to overloading, especially to higher concentrations of ethanol in the bioreactor. In Chapter 5, I examined the effect of hydrogen partial pressures on the microbiome composition and function in three chain-elongation bioreactors (5-L working volume) with product extraction. Unlike the anaerobic digestion bioreactors, these bioreactors were not found to perform similarly under similar imposed operating conditions, which indicates the instability of these systems. At the highest hydrogen partial pressure, the lowest rate of conversion of the substrate to *n*-caprylate was observed. Some limitations in the study design prevented definite conclusions from being drawn on the effect of hydrogen partial pressure in these systems. Metagenomics and metaproteomics analysis revealed no organisms present in the bioreactor with the complete reverse beta-oxidation pathway, indicating the syntrophic organisms may be involved or that the chain-elongation system is unique and reference organisms do not exist in the databases. The low percent matches to the reference database limited the conclusions that could be drawn from this analysis. In both Chapters 4 and 5, members of the *Clostridium* cluster IV within the Ruminococcaceae family, in particular *Oscillospira* species, were found to

be correlated to improved *n*-caprylate productivity. In Chapters 4 and Chapters 5, the Shannon diversity (*i.e.,* richness and evenness) of the chain-elongation bioreactor microbial communities was found to be lower than that of the anaerobic digester microbiome. However, the lower complexity of the substrate used for the chain-elongation bioreactors (*i.e.,* ethanol and acetate based medium) as compared to the cow manure substrate used in the anaerobic digestion study may be a confounding factor. It should also be noted that a lower Shannon diversity does not necessarily mean that the bioreactors did not have functionally redundant microbes present. Nevertheless, this lower Shannon diversity may have contributed to the relative instability of the chain-elongation microbiomes observed in Chapters 4 and 5.

6.2 Recommendations for Future Work

There are many avenues to explore to further expand understanding of carboxylate platform microbiomes. In Chapter 3, the study only utilized 16S rRNA gene sequencing to examine the composition of the anaerobic digester microbiome. Use of metagenomics, metaproteomics, and metatranscriptomics tools would allow future researchers to examine how monensin and other antibiotics affect the activity and function of anaerobic digester microbiomes. The studies on the chain-elongation microbiome in Chapters 4 and 5 open up many further questions as to what contributes to stability and improved performance in these systems. This work provides some evidence that shaping the carboxylate microbiome to produce medium-chain carboxylates instead of methane (*via* lower pH and product extraction) leads to a community with lower richness and evenness. Further work should seek to characterize the resistance (*i.e.,* ability of the system to resist change in response to a disturbance), resilience (*i.e.,* ability of a community to recover following a disturbance), and functional redundancy (*i.e.,* ability of other members of the

community to take over the function of negatively affected members following a disturbance) of the chain-elongation microbiome. Researchers should seek to characterize which microbes are carrying out chain elongation in these systems, whether syntrophic partnerships are involved, and to identify functionally redundant members of the community. This works points to members of the *Clostridium* cluster IV, especially *Oscillospira* members, as being possibly important players in these systems. Work should be done to isolate and characterize these microbes. In the study in Chapter 5, metagenomics and metaproteomics analysis found low matches to the reference database despite high quality indicating that many of the microbes in these systems have yet to be characterized. This work points to some paths to explore for improving the operation of these systems. These chain-elongation bioreactors in Chapters 4 and 5 were sensitive to overloading, therefore, future researchers should seek to carefully monitor the ethanol present in these bioreactor systems and determine the optimal ethanol concentrations to maintain in these systems. Longer hydraulic retention times may allow for more conversion of *n*-butyrate to *n*-caproate and then to *n*-caprylate. This work also indicates that researchers should run replicate chain-elongation bioreactors in the future to ensure that their results are repeatable. In regards to the hydrogen study in Chapter 5, future work should employ a dissolved hydrogen probe to get more accurate measurements of the hydrogen concentration experienced by the microbiome and to more fully understand the effect of hydrogen in these systems. In conclusion, the studies reported on in this dissertation suggest several promising pathways for future researchers to pursue as they seek to optimize and better characterize these systems.

APPENDIX 1

SUPPLEMENTAL INFORMATION FOR: RESISTANCE AND ADAPTATION TO THE ANTIBIOTIC MONENSIN BY THE ANAEROBIC DIGESTION MICROBIOME

Table A1.1. Anaerobic digester biomass samples. In addition to these samples, ten manure substrate samples were sequenced (duplicate samples from each manure batch), as well as the cow hindgut samples.

Day	**Period**	**Number of Samples Processed Per Anaerobic digester**			
		Low A	**Low B**	**Fast**	**Slow**
0	Startup	2	1	2	2
5	Startup	1	1	1	1
27	Startup	1	1	1	1
52	Startup	1	1	1	1
77	Startup	1	1	1	1
97	Startup	1	1	1	1
119	P1	1	1	1	1
133	P1	1	1	1	1
147	P1	1	1	1	1
175	P1	1	1	1	1
189	P1	1	1	1	1
201	P1	1	1	1	1
203	P2	1	1	1	1
209	P2	1	1	1	1
215	P2	1	1	1	1
217	P2	1	1	1	1
223	P2	1	1	1	1
231	P2	1	1	1	1
237	P2	1	1	1	1
243	P2	1	1	1	1
249	P2	1	1	1	1
255	P2	1	1	1	1
257	P2	1	1	1	1
265	P3	1	1	1	1
273	P3	1	1	1	1
281	P3	1	1	1	1
287	P3	1	1	1	1
291	P3	0	0	1	1
297	P3	1	1	1	1
299	P3	0	0	1	1

Day	Period	Number of Samples Processed Per Anaerobic digester			
		Low A	Low B	Fast	Slow
301	P3	1	1	1	1
307	P4	1	1	1	1
315	P4	1	1	0	1
321	P4	0	0	0	1
329	P4	1	1	0	1
333	P4	1	1	0	1
339	P4	0	0	0	1
343	P4	1	1	0	1
349	P4	1	1	1	1
355	P4	0	1	1	0
356	P4	0	1	1	0
357	P4	1	0	0	1
363	P4	1	0	0	1
369	P4	1	0	0	1
381	P4	1	0	0	1
383	P4	1	0	0	1

Table A1.2. Physical and chemical data from manure characterization. Values are bold when significant differences from control manure values were observed ($p<0.05$). Star (*) indicates whether the monensin manure remained significantly different after considering the dilutions of the manures. Errors are represented as standard deviations.

	Manures				
Measured parameters	**M0**	**M200**	**M300**	**M400**	**M500**
Dilution Factor	0.391	0.372	0.392	0.379	0.359
Monensin-Dosing Rate ($mg \cdot d^{-1}$)	ND	194 ± 1	320 ± 46	432 ± 19	546 ± 51
Monensin Concentration, wet matter basis ($mg \cdot L^{-1}$)	ND	0.239 ± 0.009	0.337 ± 0.008	0.371 ± 0.013	0.543 ± 0.011
Monensin Concentration, dry matter basis ($mg \cdot kg^{-1}$)	ND	1.610 ± 0.050	2.350 ± 0.170	2.510 ± 0.210	3.420 ± 0.100
pH	7.05 ± 0.22	7.19 ± 0.04	**7.58 ± 0.04***	**7.56 ± 0.11***	**7.55 ± 0.16***
Gross Energy ($cal \cdot g^{-1}$)	4590 ± 22	4607 ± 37	4583 ± 53	4543 ± 23	4612 ± 52
Total VFAs Concentration (mg $Ac \cdot L^{-1}$)	6416 ± 194	**5168 ± 125***	**5576 ± 338***	6212 ± 441	5861 ± 372*
Acetate/Propionate Ratio	7.3 ± 0.1	6.5 ± 0.5	7.0 ± 0.2	**9.0 ± 0.4***	8.0 ± 0.7
TS Concentration ($g \cdot L^{-1}$)	140 ± 1.9	**143.3 ± 4.7**	142.2 ± 1.6	**144.2 ± 2.5**	**156.1 ± 3**
VS Concentration ($g \cdot L^{-1}$)	126.8 ± 2.1	**130.0 ± 4.8**	127.0 ± 1.6	129.7 ± 2.5	**141.3 ± 3.0**
Total COD Concentration (g $O_2 \cdot L^{-1}$)	135.07 ± 19.47	146.00 ± 18.77	145.60 ± 14.35	118.40 ± 18.88	136.80 ± 13.49
Soluble COD Concentration (g $O_2 \cdot L^{-1}$)	24.40 ± 1.31	23.09 ± 0.72	26.67 ± 1.34	24.27 ± 2.36	23.20 ± 3.47
Total Alkalinity (g $CaCO_3 \cdot L^{-1}$)	9.3 ± 0.1	**10.4 ± 0.1***	**12.6 ± 0.3***	**12.5 ± 0.1***	**12.4 ± 0.1***
Total Ammonium Concentration (g $N \cdot L^{-1}$)	1.00 ± 0.04	0.90 ± 0.04	**1.44 ± 0.05***	**1.82 ± 0.18***	**1.97 ± 0.12***

Table A1.3. Relative abundance (%) of taxa families in the manure substrate samples. Only families that reached at least 1% relative abundance in any one manure substrate sample are included in the table. This resulted in 17 families that represent an average of 95.33% of the total composition of the manures. Taxa that were unclassified to the family level are included in the table (*i.e.,* Bacteroidales Unknown Family in table represents all OTUs that were classified to the order Bacteroidales but were unclassified at the family level). M0 refers to the manure from the control cows. M200, M300, M400, and M500 refer to the four batches of manure collected from the monensin-dosed cows (corresponding to the 194, 320, 432, and 546 $mg \cdot d^{-1}$ monensin-dosing rates, respectively). Duplicate or triplicate samples were run from each manure substrate and are shown in the table as separate entries.

Family Name	**Relative Abundance of Taxa Family in Manures (%)**										
	M0	**M0**	**M200**	**M200**	**M300**	**M300**	**M300**	**M400**	**M400**	**M500**	**M500**
Aerococcaceae	0.81	1.55	1.50	1.59	9.09	8.95	1.29	16.95	17.09	19.24	21.26
Bacteroidaceae	1.32	1.18	1.03	0.99	1.57	1.21	1.80	0.92	0.96	0.52	0.77
Bacteroidales Family S24-7	4.48	3.68	2.76	2.65	3.09	2.44	3.22	1.55	1.67	1.61	1.70
Bacteroidales Unknown Family	8.35	10.81	4.76	3.97	8.09	5.73	7.60	1.99	2.45	2.57	3.33
Bifidobacteriaceae	5.67	5.12	0.10	0.10	0.06	0.07	0.06	1.47	1.11	0.84	0.88
Carnobacteriaceae	0.61	3.60	1.54	1.81	13.20	21.67	1.74	19.48	19.59	19.63	15.14
Clostridiaceae	4.85	4.17	7.82	8.53	3.57	2.82	5.77	3.03	2.90	5.52	6.25
Clostridiales Unknown Family	10.07	9.17	17.45	17.70	9.40	8.05	11.96	8.43	8.47	6.74	6.74
Erysipelotrichaceae	1.22	0.73	0.49	0.36	0.81	0.58	0.75	0.56	0.49	0.37	0.38
Lachnospiraceae	7.50	6.85	6.48	7.84	6.49	5.30	8.78	6.07	5.89	7.04	5.64
Methanobacteriaceae	1.73	1.76	0.77	0.62	1.04	0.74	0.71	0.81	0.59	1.12	1.10
Mollicutes RF39 Order, Unknown Family	5.95	6.20	10.15	9.12	5.49	5.24	5.49	4.84	4.88	3.08	3.52
Moraxellaceae	0.37	0.32	0.25	0.24	0.08	0.08	0.10	0.44	0.54	1.32	0.84
Rikenellaceae	0.99	0.83	0.96	0.89	1.36	1.15	1.40	0.49	0.59	0.37	0.55
Ruminococcaceae	41.17	39.55	38.74	38.48	30.01	30.21	43.55	27.47	27.76	24.25	25.57
Staphylococcaceae	0.03	0.07	0.28	0.29	0.24	0.25	0.12	0.67	0.54	0.69	1.25
Turicibacteraceae	0.45	0.44	0.08	0.09	1.26	0.90	0.89	0.17	0.11	0.14	0.16

Table A1.4. iVFAs and tVFA concentrations. The former was measured with gas chromatography and the latter was measured with the total VFAs measurement method.

Anaerobic digester	Day	Acetate concentration (mM)	Propionate concentration (mM)	*n*-Butyrate concentration (mM)	tVFA concentration measured (mmol Ac·L^{-1})
Low A	219	0.24	ND	ND	1.69
Low A	245	0.20	ND	ND	1.69
Low A	285	0.14	ND	ND	1.69
Low A	345	0.11	ND	ND	3.17
Low B	247	0.13	ND	ND	1.79
Low B	323	0.17	ND	ND	2.34
Low B	333	0.11	ND	ND	2.34
Low B	345	0.42	ND	ND	3.01
Low B	353	0.90	ND	ND	3.39
Fast	243	0.62	0.45	ND	1.60
Fast	285	7.36	ND	ND	4.35
Fast	293	38.98	3.69	1.15	55.57
Fast	307	38.61	9.86	2.81	102.69
Slow	321	0.52	ND	ND	2.18
Slow	323	1.31	ND	ND	3.33
Slow	333	11.64	1.68	ND	28.61
Slow	345	0.56	ND	ND	3.33
Slow	353	0.38	ND	ND	2.72

Table A1.5. Stability parameters measured for each anaerobic digester. Errors are represented as standard deviations.

Period	**Startup (Days 1-114)**				**Period 1 (Days 115-202)**			
Anaerobic digester	LowA	LowB	Fast	Slow	LowA	LowB	Fast	Slow
Stability Parameters								
pH	7.37 ± 0.07	7.38 ± 0.08	7.39 ± 0.08	7.38 ± 0.07	7.23 ± 0.09	7.23 ± 0.09	7.22 ± 0.08	7.23 ± 0.11
tVFA Concentration (mg $Ac \cdot L^{-1}$)	184 ± 38	250 ± 123	237 ± 173	165 ± 39	161 ± 23	154 ± 21	130 ± 14	140 ± 28
Total Alkalinity (g $CaCO_3 \cdot L^{-1}$)	NA	NA	NA	NA	5.7 ± 0.1	5.8 ± 0.1	5.8 ± 0.1	5.9 ± 0.1
Bicarbonate Alkalinity (g $CaCO_3 \cdot L^{-1}$)	NA	NA	NA	NA	5.6 ± 0.1	5.7 ± 0.1	5.7 ± 0.1	5.8 ± 0
TS Concentration ($g \cdot L^{-1}$)	38.3 ± 3.2	39.1 ± 3.9	39.7 ± 3.3	39.2 ± 2.8	31.5 ± 2.4	31.7 ± 2.3	32.2 ± 1.2	30.7 ± 2.2
VS Concentration ($g \cdot L^{-1}$)	28.3 ± 2.2	28.7 ± 2.3	29 ± 2.1	28.7 ± 1.7	26.1 ± 2	26.2 ± 2.1	26.9 ± 1	25.3 ± 1.9
Soluble COD concentration (g $O_2 \cdot L^{-1}$)	NA	NA	NA	NA	3.48 ± 0.66	4.43 ± 2.09	3.33 ± 0.63	3.72 ± 0.98
Total Ammonium concentration (g $N \cdot L^{-1}$)	NA	NA	NA	NA	0.79 ± 0.01	0.81 ± 0.02	0.75 ± 0.05	0.84 ± 0.02
Period	**Period 2 (Days 203-264)**				**Period 3 (Days 265-306)**			
Anaerobic digester	LowA	LowB	Fast	Slow	LowA	LowB	Fast	Slow
Stability Parameters								
pH	7.22 ± 0.07	7.22 ± 0.08	7.17 ± 0.08	7.21 ± 0.09	7.22 ± 0.04	7.21 ± 0.05	6.71 ± 0.48	7.14 ± 0.07
tVFA Concentration (mg $Ac \cdot L^{-1}$)	117 ± 23	119 ± 18	112 ± 23	107 ± 20	128 ± 15	122 ± 17	2473 ± 2426	108 ± 10
Total Alkalinity (g	6.4 ± 0.5	6.3 ± 0.6	5.9 ± 0.1	5.9 ± 0.1	7 ± 0.1	6.9 ± 0.1	5.7 ± 0.4	6 ± 0.1

$CaCO_3 \cdot L^{-1}$)								
Bicarbonate Alkalinity (g $CaCO_3 \cdot L^{-1}$)	6.3 ± 0.5	6.2 ± 0.6	5.8 ± 0.2	5.8 ± 0.1	6.9 ± 0.1	6.8 ± 0.1	3.8 ± 2.1	6 ± 0.1
TS Concentration ($g \cdot L^{-1}$)	31 ± 2.5	32.1 ± 1.9	31.5 ± 2.3	31.1 ± 0.8	32.9 ± 2.7	31.3 ± 2.2	34.1 ± 3.1	30.9 ± 1.2
VS Concentration ($g \cdot L^{-1}$)	26.1 ± 2	26.8 ± 2	26.5 ± 2.3	26.3 ± 0.8	27.7 ± 2.5	26 ± 2	29 ± 3	25.9 ± 1.4
Soluble COD concentration (g $O_2 \cdot L^{-1}$)	3.29 ± 0.47	3.27 ± 0.3	3.09 ± 0.32	3.28 ± 0.45	3.63 ± 0.28	3.33 ± 0.34	7.34 ± 5.11	3.28 ± 0.5
Total Ammonium concentration (g $N \cdot L^{-1}$)	0.95 ± 0.11	0.96 ± 0.11	0.85 ± 0.05	0.82 ± 0.04	1.11 ± 0.1	1.12 ± 0.08	0.96 ± 0.1	0.84 ± 0.05
Period	**Period 4 (Days 307-383)**							
Anaerobic digester	LowA	LowB	Fast	Slow				
Stability Parameters								
pH	7.23 ± 0.08	7.22 ± 0.07		7.14 ± 0.1				
tVFA Concentration (mg $Ac \cdot L^{-1}$)	155 ± 29	162 ± 27		384 ± 446				
Total Alkalinity (g $CaCO_3 \cdot L^{-1}$)	7.2 ± 0.2	7.2 ± 0.1		6.3 ± 0.2				
Bicarbonate Alkalinity (g $CaCO_3 \cdot L^{-1}$)	7 ± 0.2	7.1 ± 0.1		6 ± 0.3				
TS Concentration ($g \cdot L^{-1}$)	35 ± 1.9	32.4 ± 1.2		34.1 ± 1.9				
VS Concentration ($g \cdot L^{-1}$)	29.8 ± 1.9	27.1 ± 1.1		29.2 ± 1.8				
Soluble COD concentration (g $O_2 \cdot L^{-1}$)	4.13 ± 0.38	3.77 ± 0.39		4.13 ± 0.69				
Total Ammonium concentration (g $N \cdot L^{-1}$)	1.13 ± 0.1	1.2 ± 0.14		0.9 ± 0.15				

Table A1.6. Performance parameters measured for each anaerobic digester. Errors are represented as standard deviations.

Period	**Startup (Days 1-114)**				**Period 1 (Days 115-202)**			
Anaerobic digester	LowA	LowB	Fast	Slow	LowA	LowB	Fast	Slow
Performance								
Biogas Production Rate ($L \cdot L^{-1} \cdot d^{-1}$)	0.453 ± 0.061	0.467 ± 0.071	0.356 ± 0.124	0.444 ± 0.071	0.741 ± 0.072	0.76 ± 0.06	0.727 ± 0.056	0.718 ± 0.042
Specific Methane Yield ($L\ CH_4 \cdot gVS^{-1}d^{-1}$)	NA	NA	NA	NA	0.21 ± 0.014	0.211 ± 0.018	0.188 ± 0.017	0.195 ± 0.013
Methane Content (%)	NA	NA	NA	NA	54 ± 1	54 ± 1	54 ± 1	54 ± 1
VS Reduction Efficiency (%)	43.4 ± 4.4	42.7 ± 4.6	42 ± 4.2	42.7 ± 3.5	47.7 ± 4.1	47.6 ± 4.3	46.3 ± 2	49.5 ± 3.8
Soluble COD Reduction Efficiency (%)	NA	NA	NA	NA	63 ± 7	54 ± 22	65 ± 7	61 ± 11
Period	**Period 2 (Days 203-264)**				**Period 3 (Days 265-306)**			
Anaerobic digester	LowA	LowB	Fast	Slow	LowA	LowB	Fast	Slow
Performance								
Biogas Production Rate ($L \cdot L^{-1} \cdot d^{-1}$)	0.706 ± 0.11	0.716 ± 0.105	0.755 ± 0.043	0.732 ± 0.046	0.608 ± 0.032	0.653 ± 0.052	0.409 ± 0.275	0.722 ± 0.044
Specific Methane Yield ($L\ CH_4 \cdot gVS^{-1}d^{-1}$)	0.189 ± 0.027	0.192 ± 0.026	0.2 ± 0.013	0.195 ± 0.014	0.17 ± 0.009	0.18 ± 0.015	0.108 ± 0.086	0.2 ± 0.013
Methane Content (%)	54 ± 3	54 ± 3	53 ± 3	53 ± 2	56 ± 2	55 ± 1	44 ± 16	55 ± 2
VS Reduction Efficiency (%)	47.7 ± 4	46.4 ± 3.9	47 ± 4.5	47.4 ± 1.6	44.8 ± 5.1	48.1 ± 4	42.1 ± 6	48.1 ± 2.6
Soluble COD Reduction Efficiency (%)	64 ± 7	65 ± 5	67 ± 3	66 ± 5	57 ± 3	60 ± 4	37 ± 35	65 ± 5
Period	**Period 4 (Days 307-383)**							

Anaerobic digester	LowA	LowB	Fast	Slow	
Performance					
Biogas Production Rate ($L \cdot L^{-1} \cdot d^{-1}$)	0.587 ± 0.033	0.636 ± 0.034		0.587 ± 0.075	
Specific Methane Yield (L $CH_4 \cdot gVS^{-1}d^{-1}$)	0.161 ± 0.009	0.176 ± 0.009		0.159 ± 0.022	
Methane Content (%)	55 ± 1	55 ± 1		55 ± 2	
VS Reduction Efficiency (%)	40.4 ± 3.7	42.1 ± 6		41.5 ± 3.6	
Soluble COD Reduction Efficiency (%)	50 ± 5	55 ± 5		57 ± 7	

Table A1.7. Average relative abundance (%) of taxa families in the seven digester inoculum (Day 0) samples. Only families that reached at least 1% relative abundance in any one digester inoculum sample are included in this table. Taxa that were unclassified to the family level are included in the table (*i.e.,* Bacteroidales Unknown Family in table represents all OTUs that were classified to the order Bacteroidales but were unclassified at the family level). This resulted in 20 families that represent 86.10% of the total composition of the inoculum samples.

Taxa Family	**Percent Average Relative Abundance (mean±sd)**
Anaerolinaceae	6.27±0.69
Bacteroidales Family SB-1	0.96±0.09
Bacteroidales Unknown Family	26.00±1.58
Cloacamonaceae	9.23±1.26
Clostridia order SHA-98, Unknown Family	2.69±0.43
Clostridia Unknown Family	1.01±0.10
Comamonadaceae	1.79±0.26
Geobacteraceae	2.66±0.40
Methanobacteriaceae	7.86±1.33
Methanosaetaceae	1.41±0.29
Moraxellaceae	2.50±1.66
OP8 Candidate Phyllum, Class OP8_2, Unknown Family	1.12±0.18
Pseudomonadaceae	3.28±1.42
Spirochaetaceae	5.60±1.53
Syntrophaceae	1.57±0.20
Syntrophobacteraceae	1.59±0.26
Syntrophomonadaceae	0.91±0.10
Tenericutes, class RF3, order ML615J-28, Unknown Family	1.70±0.27
Thermotogaceae	2.96±0.64
Tissierellaceae	4.98±0.96

Table A1.8. OTUs positively correlated (Spearman's rank correlation, ρ >0.5, p<0.001) with monensin concentrations in the substrate for the Slow anaerobic digester from day 175 on. Table also indicates if the OTU was positively correlated with monensin concentrations in the Fast anaerobic digester (Y=yes; N=no) and if it was positively correlated with monensin concentrations in all anaerobic digesters (*i.e.*, Low A, Low B, Fast, and Slow). Only OTUs that reached a min of 1% relative abundance in at least one anaerobic digester sample (from day 175 on) are included.

OTU ID	Taxonomy	Fast Digester	All Digesters
842598	k__Archaea; p__Euryarchaeota; c__Methanobacteria; o__Methanobacteriales; f__Methanobacteriaceae; g__Methanobrevibacter; s__	N	N
628811	k__Bacteria; p__Armatimonadetes; c__SJA-176; o__TP122; f__; g__; s__	N	N
837605	k__Bacteria; p__Bacteroidetes; c__Bacteroidia; o__Bacteroidales; f__; g__; s__	Y	Y
559410	k__Bacteria; p__Bacteroidetes; c__Bacteroidia; o__Bacteroidales; f__; g__; s__	N	N
4059526	k__Bacteria; p__Bacteroidetes; c__Bacteroidia; o__Bacteroidales; f__; g__; s__	N	N
4443296	k__Bacteria; p__Bacteroidetes; c__Bacteroidia; o__Bacteroidales; f__; g__; s__	N	N
1891861	k__Bacteria; p__Bacteroidetes; c__Bacteroidia; o__Bacteroidales; f__[Paraprevotellaceae]; g__YRC22; s__	N	N
541252	k__Bacteria; p__Bacteroidetes; c__Bacteroidia; o__Bacteroidales; f__Porphyromonadaceae; g__; s__	Y	Y
3506234	k__Bacteria; p__Bacteroidetes; c__Bacteroidia; o__Bacteroidales; f__Porphyromonadaceae; g__; s__	Y	Y
4383641	k__Bacteria; p__Bacteroidetes; c__Bacteroidia; o__Bacteroidales; f__Porphyromonadaceae; g__; s__	Y	Y
1764554	k__Bacteria; p__Bacteroidetes; c__Bacteroidia; o__Bacteroidales; f__Porphyromonadaceae; g__Parabacteroides; s__	N	N
258334	k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Anaerolineales; f__Anaerolinaceae; g__T78; s__	Y	N
809315	k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Anaerolineales; f__Anaerolinaceae; g__T78; s__	Y	N
591951	k__Bacteria; p__Firmicutes; c__Clostridia; o__Clostridiales; f__; g__; s__	N	N
178657	k__Bacteria; p__Firmicutes; c__Clostridia; o__Clostridiales; f__Clostridiaceae; g__Clostridium; s__	N	N
675613	k__Bacteria; p__Firmicutes; c__Clostridia; o__Clostridiales; f__Peptococcaceae;	Y	Y

OTU ID	Taxonomy	Fast Digester	All Digesters
	g__Pelotomaculum; s__		
550485	k__Bacteria; p__Firmicutes; c__Clostridia; o__Clostridiales; f__Ruminococcaceae; g__; s__	N	N
582502	k__Bacteria; p__Firmicutes; c__Clostridia; o__Clostridiales; f__Ruminococcaceae; g__; s__	N	N
743075	k__Bacteria; p__Firmicutes; c__Clostridia; o__Clostridiales; f__Ruminococcaceae; g__; s__	N	N
539716	k__Bacteria; p__Firmicutes; c__Clostridia; o__Clostridiales; f__Syntrophomonadaceae; g__Syntrophomonas; s__	N	N
820298	k__Bacteria; p__OP11; c__WCHB1-64; o__d153; f__; g__; s__	N	N
286595	k__Bacteria; p__Proteobacteria; c__Deltaproteobacteria; o__Desulfovibrionales; f__Desulfomicrobiaceae; g__Desulfomicrobium; s__	N	N
4353504	k__Bacteria; p__Spirochaetes; c__Spirochaetes; o__Spirochaetales; f__Spirochaetaceae; g__Treponema; s__	N	N
767422	k__Bacteria; p__Synergistetes; c__Synergistia; o__Synergistales; f__Dethiosulfovibrionaceae; g__HA73; s__	Y	Y
703270	k__Bacteria; p__Synergistetes; c__Synergistia; o__Synergistales; f__Synergistaceae; g__vadinCA02; s__	N	N
616150	k__Bacteria; p__Verrucomicrobia; c__Opitutae; o__Puniceicoccales; f__Puniceicoccaceae; g__; s__	N	N

Table A1.9. OTUs negatively correlated (Spearman's rank correlation, ρ <-0.5, p<0.001) with monensin concentrations in substrate for the Slow anaerobic digester from day 175 on. Table also indicates if the OTU was negatively correlated with monensin concentrations in the Fast anaerobic digester (Y=yes; N=no) and if it was negatively correlated with monensin concentrations in all anaerobic digesters (*i.e.*, Low A, Low B, Fast, and Slow). Only OTUs that reached a min of 1% relative abundance in at least one anaerobic digester sample (from day 175 on) are included.

OTU ID	Taxonomy	Fast Digester	All Digesters
752382	k__Bacteria; p__Bacteroidetes; c__Bacteroidia; o__Bacteroidales; f__; g__; s__	Y	Y
111350	k__Bacteria; p__Bacteroidetes; c__Bacteroidia; o__Bacteroidales; f__; g__; s__	Y	N
1144358	k__Bacteria; p__Bacteroidetes; c__Bacteroidia; o__Bacteroidales; f__; g__; s__	Y	N

OTU ID	Taxonomy	Fast Digester	All Digesters
4471301	k__Bacteria; p__Bacteroidetes; c__Bacteroidia; o__Bacteroidales; f__; g__; s__	Y	N
509055	k__Bacteria; p__Bacteroidetes; c__Bacteroidia; o__Bacteroidales; f__; g__; s__	N	N
562538	k__Bacteria; p__Bacteroidetes; c__Bacteroidia; o__Bacteroidales; f__; g__; s__	N	N
681365	k__Bacteria; p__Bacteroidetes; c__Bacteroidia; o__Bacteroidales; f__; g__; s__	N	N
787023	k__Bacteria; p__Bacteroidetes; c__Bacteroidia; o__Bacteroidales; f__; g__; s__	N	N
4468051	k__Bacteria; p__Bacteroidetes; c__Bacteroidia; o__Bacteroidales; f__Bacteroidaceae; g__; s__	Y	N
549570	k__Bacteria; p__Bacteroidetes; c__Bacteroidia; o__Bacteroidales; f__SB-1; g__; s__	Y	Y
550996	k__Bacteria; p__Firmicutes; c__Clostridia; o__Clostridiales; f__Lachnospiraceae; g__Coprococcus; s__	Y	N
3273469	k__Bacteria; p__Firmicutes; c__Clostridia; o__Clostridiales; f__Ruminococcaceae; g__; s__	Y	Y
562603	k__Bacteria; p__Proteobacteria; c__Deltaproteobacteria; o__Syntrophobacterales; f__Syntrophaceae; g__Syntrophus; s__	Y	Y
736489	k__Bacteria; p__Proteobacteria; c__Deltaproteobacteria; o__Syntrophobacterales; f__Syntrophaceae; g__Syntrophus; s__	Y	Y
279151	k__Bacteria; p__Tenericutes; c__Mollicutes; o__RF39; f__; g__; s__	Y	Y
111131	k__Bacteria; p__WWE1; c__[Cloacamonae]; o__[Cloacamonales]; f__[Cloacamonaceae]; g__Candidatus Cloacamonas; s__	Y	Y
534698	k__Bacteria; p__WWE1; c__[Cloacamonae]; o__[Cloacamonales]; f__[Cloacamonaceae]; g__W5; s__	N	N

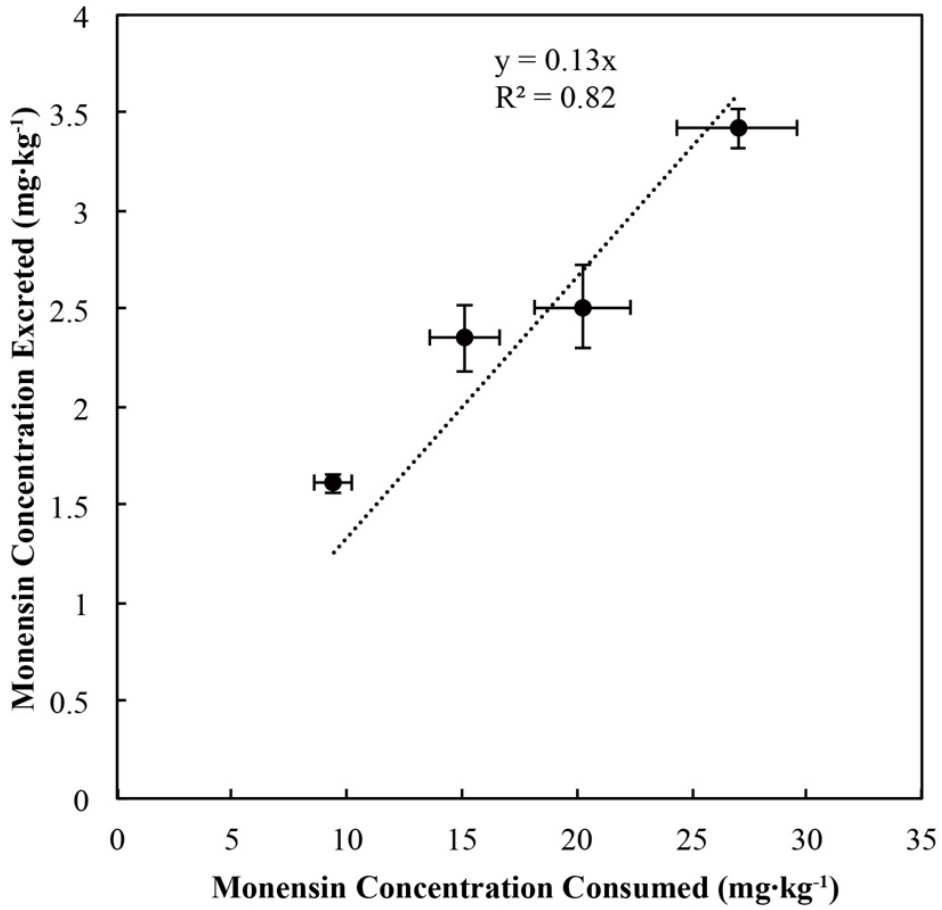

Figure A1.1. Monensin concentration measured in manure *vs.* monensin concentration measured in the consumed cow feed (on a dry matter basis in $mg \cdot kg^{-1}$). The monensin excretion rate for the cows was calculated as the slope of the best-fit line for the plot.

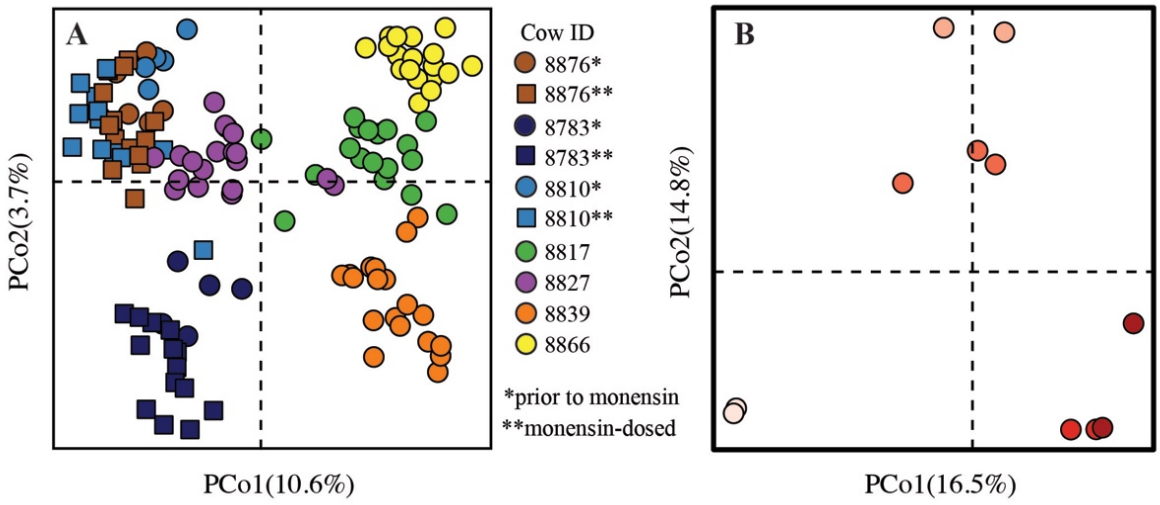

Figure A1.2. β-diversity of gut microbiome samples from cow hindguts from the control cows and monensin-dosed cows: **A**) PCoA based on unweighted Unifrac distance. The samples were collected starting in the two weeks *prior* to monensin dosing through to end of the monensin dosing period (three months total). The samples are colored based on which cow they were collected from. Square shapes represent samples collected when the cows were dosed with monensin, while circles represent time periods when the cows were not fed monensin; and **B**) PCoA of unweighted Unifrac distance of combined manure substrates that were fed to anaerobic digesters. Red shading represents the concentration of monensin measured in the manure (M0, M200, M300, M400, and M500).

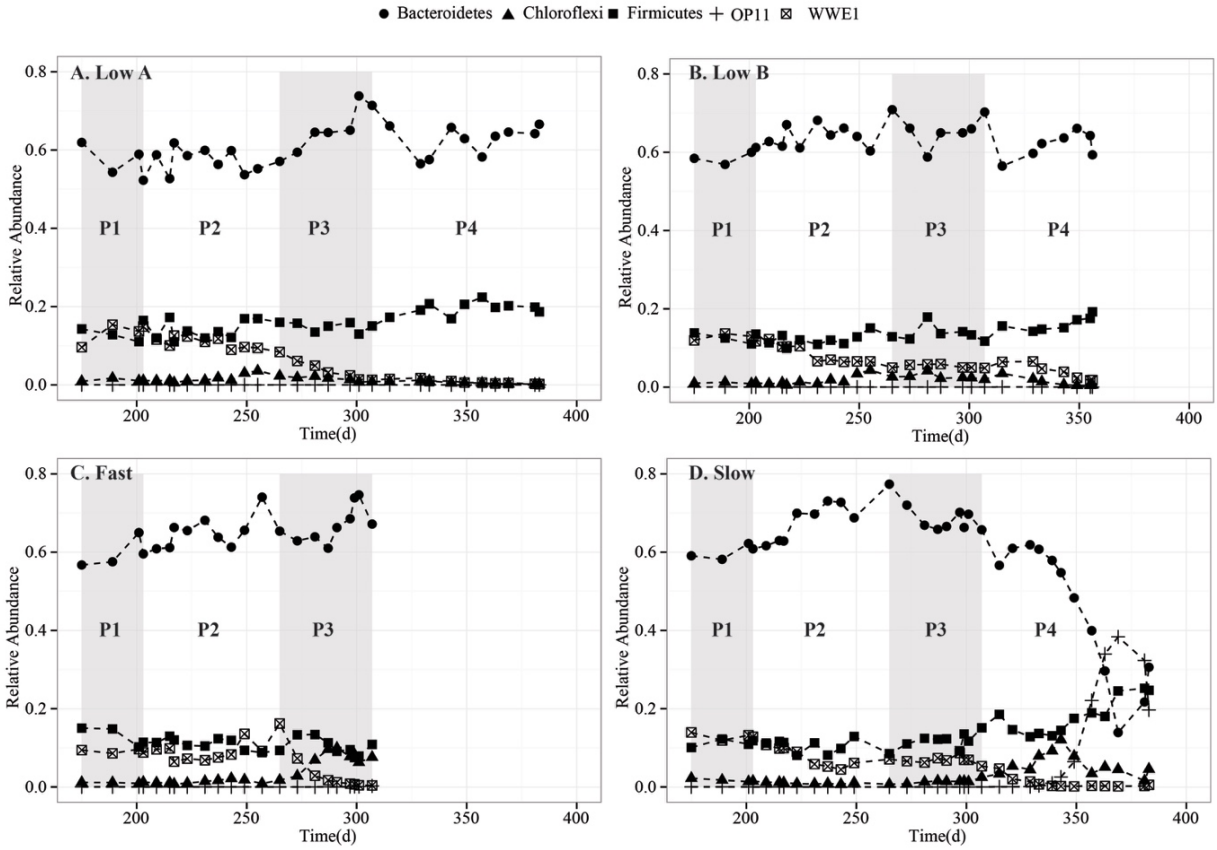

Figure A1.3. Plot of phyla that reached over 10% relative abundance in any one anaerobic digester sample from Day 175 to the end of the operating period for each of the anaerobic digesters: **A**) Low A; **B**) Low B; **C**) Fast; and **D**) Slow anaerobic digester. Shading represents periods.

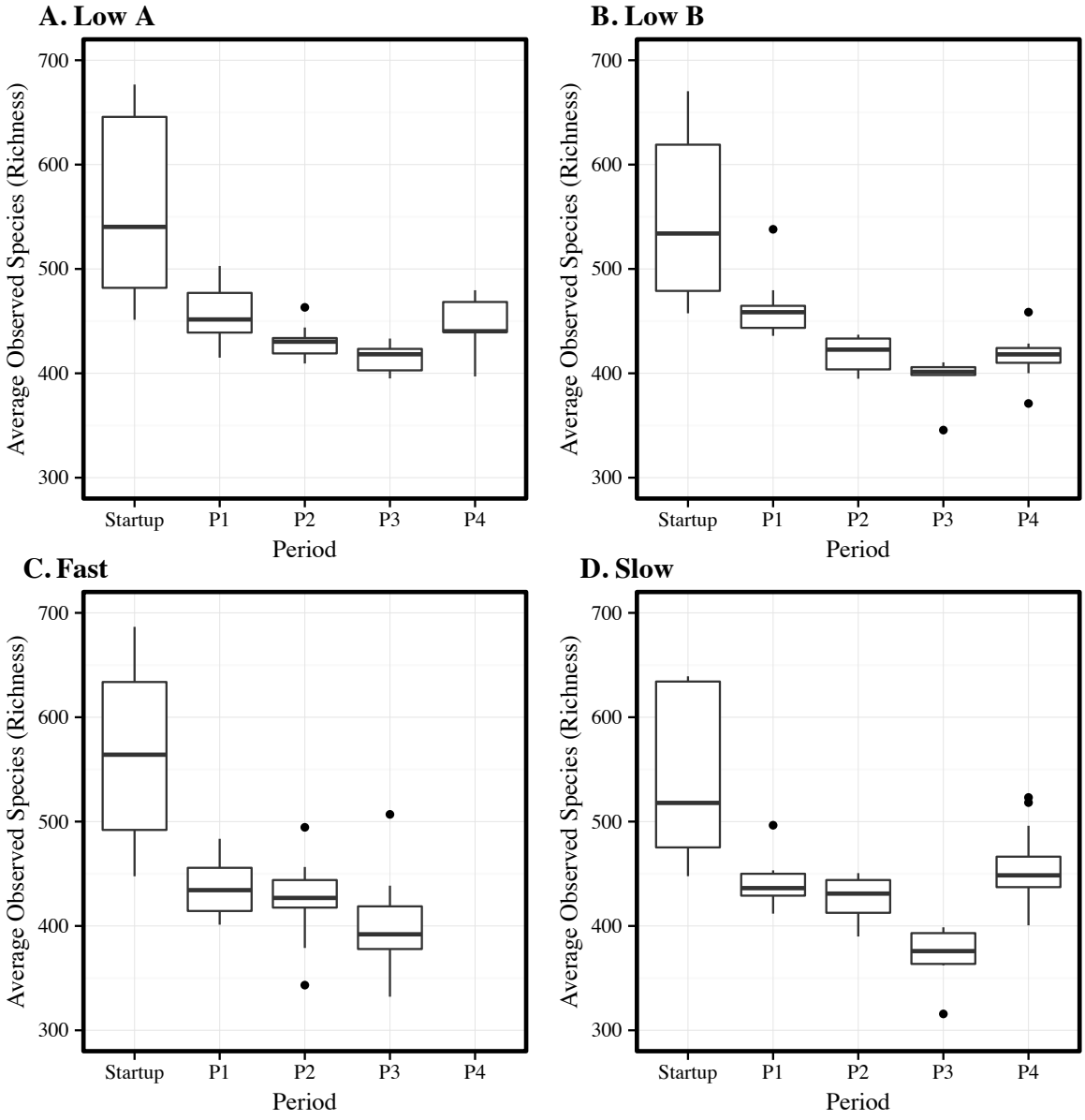

Figure A1.4. Averaged observed OTUs (richness) metric during different periods for each of the anaerobic digesters: **A**) Low A; **B**) Low B; **C**) Fast; and **D**) Slow anaerobic digester.

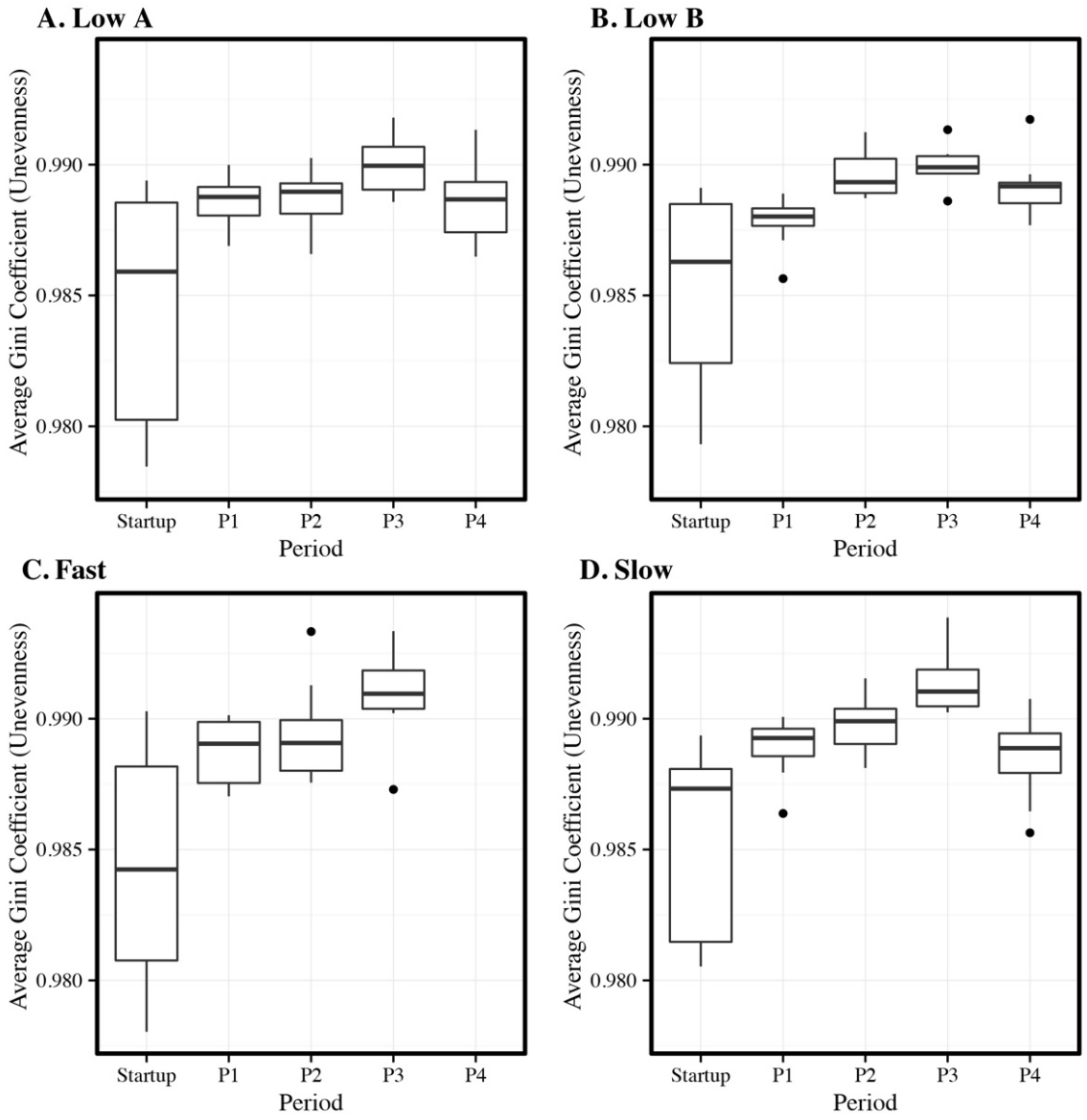

Figure A1.5. Averaged Gini coefficient (unevenness) metric during different periods for each of the anaerobic digesters: **A**) Low A; **B**) Low B; **C**) Fast; and **D**) Slow anaerobic digester.

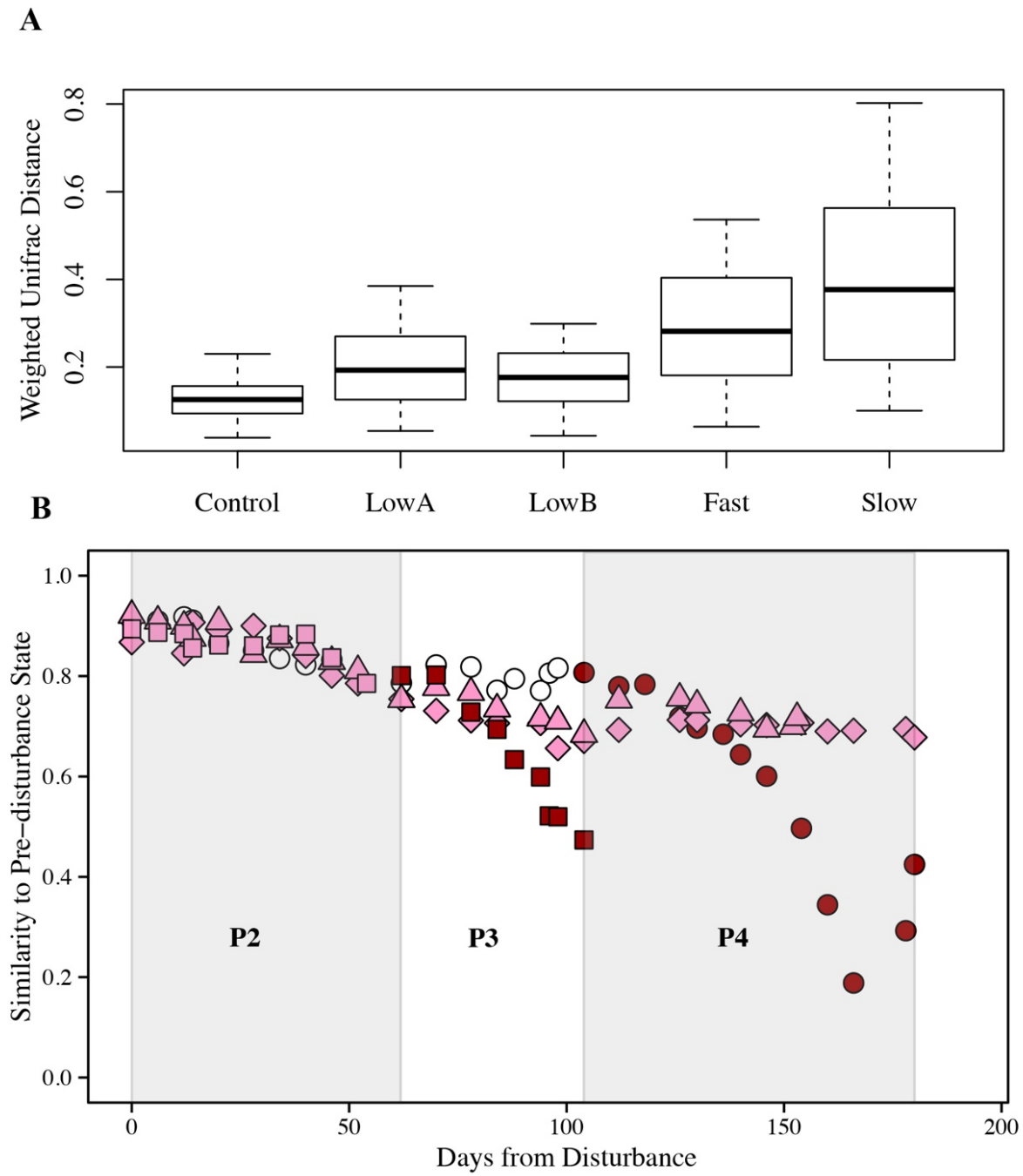

Figure A1.6. β-diversity of anaerobic digester microbiome samples during different periods: **A**) boxplot of weighted Unifrac distance between anaerobic digester samples. Control anaerobic digester boxplot corresponds to days 203 to 306 when the Slow anaerobic digester was fed only control cow manure; Low A anaerobic digester boxplot corresponds to days 203 to 383; Low B anaerobic digester boxplot corresponds to days 203 to 355, the Fast anaerobic digester boxplot only includes samples during which the anaerobic digester was subjected to high concentrations of monensin (*i.e.,* days 265 to 306), and the Slow anaerobic digester boxplot corresponds to days

306 to 383 when the anaerobic digester was subjected to high concentrations of monensin; and **B**) similarity of samples post-disturbance to samples *prior* to disturbance (*prior* to Period 2) based on the weighted Unifrac distance metric. Similarity was calculated as one minus the average weighted Unifrac distance between the sample day and the three sample days *prior* to disturbance (*prior* to period 2). The symbols represent the different anaerobic digesters: Diamond=LowA anaerobic digester; Triangle=LowB anaerobic digester; Square=Fast anaerobic digester; and Circle=Slow anaerobic digester. Color represents the monensin concentrations gradient: Dark red=high monensin concentration (1 to 5 $mg \cdot L^{-1}$), pink=low monensin concentration (<1 $mg \cdot L^{-1}$), white=no monensin in substrate. Periods (P2 to P4) are represented by gray and white shading.

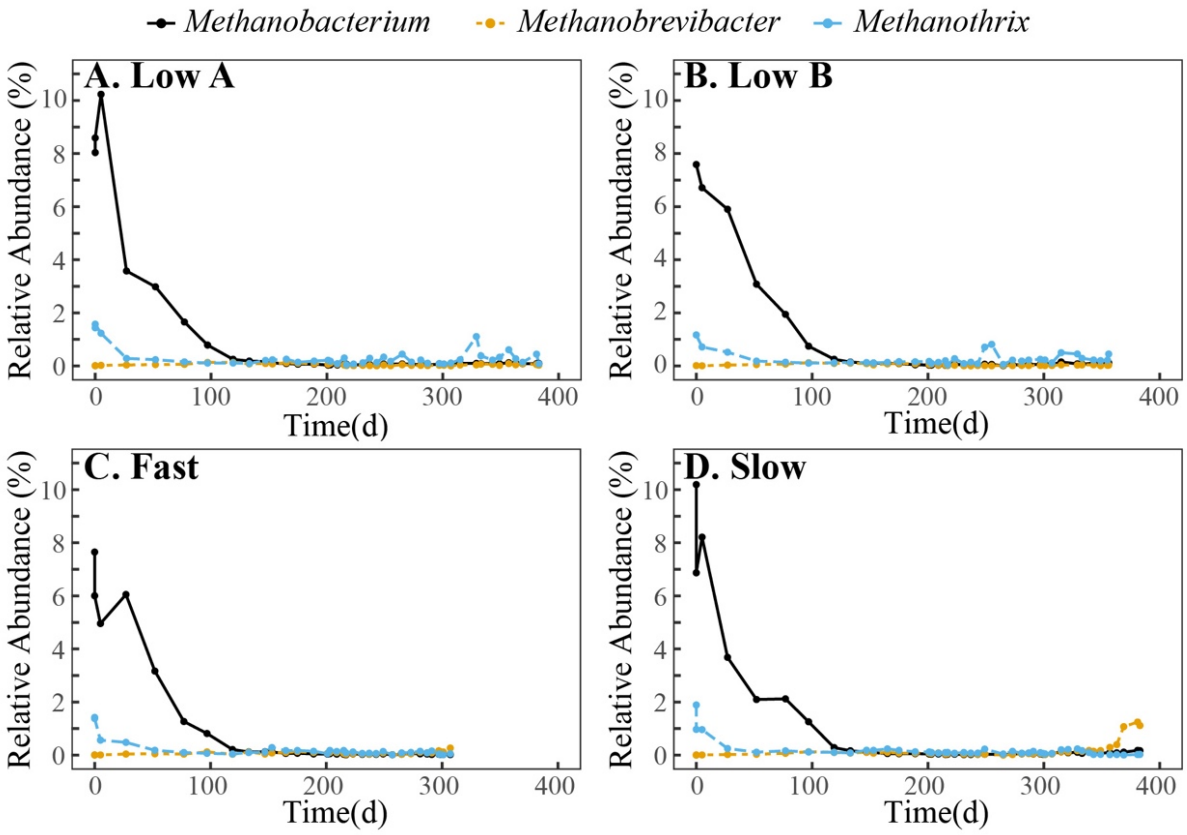

Figure A1.7. Relative abundance (%) of three methanogenic genera (*Methanobacterium*, *Methanobrevibacter*, and *Methanothrix*) during the operating period for the four anaerobic digesters: **A**) Low A; **B**) Low B; **C**) Fast; and **D**) Slow anaerobic digester. These three genera were the only methanogenic genera that reached over 1% relative abundance in any one digester sample over the time-series.

APPENDIX 2

SUPPLEMENTAL INFORMATION FOR: HIGHER SUBSTRATE RATIOS OF ETHANOL TO ACETATE DRIVE PRODUCTION TOWARD *N*-CAPRYLATE IN BIOREACTOR WITH PRODUCT EXTRACTION

Table A2.1. Specificities *per* period (*n*-butyrate, *n*-caproate, or *n*-caprylate production rate *vs*. total production rate of these carboxylates, on a COD basis) and C8 to C6 productivity ratio (on a COD basis). The substrate COD ratio (ethanol-to-acetate) is also shown for reference. Mean ± standard error is reported.

			Specificity (%)		
Period	**Substrate COD Ratio**	**C8 to C6 Ratio**	***n*-Butyrate**	***n*-Caproate**	***n*-Caprylate**
Period 1	11.7 ± 1.0	3.0 ± 0.3	3.57 ± 0.4	23.85 ± 2.05	72.57 ± 8.2
Period 2	274.9 ± 58.3	5.9 ± 0.7	4.21 ± 0.98	13.85 ± 1.57	81.94 ± 10.17
Period 3	16.9 ± 2.1	2.0 ± 0.4	3.7 ± 0.91	32.64 ± 5.45	63.66 ± 14.03
Period 4	6.7 ± 0.4	0.7 ± 0.1	14.2 ± 1.54	51.34 ± 3.33	34.47 ± 3.61
Period 5	3.6 ± 0.3	0.6 ± 0.2	33.38 ± 3.81	40.97 ± 6.41	25.65 ± 5.96
Period 6	2.9 ± 0.1	0.7 ± 0.1	40.51 ± 3.89	35.8 ± 4.5	23.69 ± 4.31
Period 7	1.8 ± 0.1	NA	98.31 ± 22.17	1.69 ± 0.28	NA
Period 8	5.1 ± 0.5	1.2 ± 0.2	64.44 ± 3.71	15.93 ± 1.75	19.62 ± 3.12
Period 9	6.5 ± 0.5	0.1 ± 0.0	92.97 ± 5.72	6.25 ± 0.47	0.78 ± 0.07
Period 10	6*	1.0 ± 0.3	59.24 ± 12.78	20.34 ± 2.51	20.42 ± 6.86
Period 11	135.2 ± 18.1	0.5 ± 0.2	56.09 ± 7.56	29.99 ± 6.41	13.92 ± 6.14

Table A2.2. Average effluent concentrations (mM) in bioreactor *per* period (mean±s.e.). nd if concentration was below detection limit (0.2 mM). If only one measurement was above detection limit, only that measurement is reported. Valerate, isocaproate, and *n*-heptanoate were not detected in the bioreactor during the study. In this table period names are abbreviated as P1 to P11 representing Periods 1 to 11. Mean ± standard error is reported. n.d. indicates that compound was below detection limit of GC (*i.e.,* 0.2 mM).

	Concentration in Effluent (mM)							
Period	**Ethanol**	**Acetate**	**Propionate**	**Isobutyrate**	***n*-Butyrate**	**Isovalerate**	***n*-Caproate**	***n*-Caprylate**
P1	15.6 ± 1	5.1 ± 0.3	nd	0.4 ± 0	4.7 ± 0.4	nd	4.5 ± 0.2	1.5 ± 0.1
P2	67.3 ± 8.8	4.1 ± 0.6	nd	0.4 ± 0	3.6 ± 0.9	0.4	2.9 ± 0.4	1.4 ± 0.1
P3	88.3 ± 1.9	12.6 ± 1.5	nd	nd	2.5 ± 0.2	nd	2.8 ± 0.2	0.9 ± 0.1
P4	49.7 ± 4.8	16 ± 1.5	nd	0.4 ± 0	9 ± 1.3	nd	2.5 ± 0.2	0.3 ± 0.1
P5	48.5 ± 4.3	35.6 ± 2.2	nd	0.5 ± 0	18.1 ± 0.6	nd	0.9 ± 0.1	nd
P6	54.8 ± 4	56.3 ± 4.9	nd	nd	20 ± 1.2	nd	0.5 ± 0	nd
P7	49.7 ± 3.2	88.2 ± 8.4	nd	0.4	18.3 ± 0.9	nd	0.4 ± 0.0	0.2
P8	58.5 ± 5.3	28.6 ± 5.4	0.5 ± 0	0.7 ± 0	28 ± 1.4	nd	0.5 ± 0	0.2
P9	98.6 ± 11.5	32.6 ± 2.7	0.6 ± 0	1.8 ± 0.1	59.7 ± 1.6	0.7 ± 0	2.5 ± 0.1	0.2 ± 0
P10	113.1 ± 4.6	41.5 ± 3.9	nd	1.2 ± 0.1	31.2 ± 2.3	nd	0.3 ± 0	nd
P11	157 ± 12.5	6.4 ± 1	nd	0.6 ± 0.1	8.6 ± 0.4	nd	0.6 ± 0.2	0.2 ± 0

Table A2.3. Conversion efficiencies (individual carboxylate production rate divided by organic loading rate) for *n*-butyrate, *n*-caproate, *n*-caprylate, and the combined medium-chain carboxylic acids (MCCA; includes *n*-caproate and *n*-caprylate) *per* period. Substrate COD ratio (ethanol-to-acetate) is provided for reference. In this table, periods are abbreviated as P1 to P11 representing Periods 1 to 11. Mean ± standard error is reported.

		Conversion Efficiency (%)			
Period	**Substrate COD Ratio**	***n*-Butyrate**	***n*-Caproate**	***n*-Caprylate**	**MCCA**
P1	11.7 ± 1.0	3.36 ± 0.33	22.46 ± 1.55	68.34 ± 6.87	90.8 ± 7.44
P2	274.9 ± 58.3	2.21 ± 0.49	7.29 ± 0.67	43.11 ± 4.52	50.4 ± 4.7
P3	16.9 ± 2.1	2.24 ± 0.5	19.75 ± 2.53	38.53 ± 7.42	58.28 ± 8.19
P4	6.7 ± 0.4	9.75 ± 1.06	35.25 ± 2.3	23.67 ± 2.49	58.92 ± 3.84
P5	3.6 ± 0.3	23.36 ± 2.15	28.67 ± 4.05	17.95 ± 3.99	46.62 ± 5.88
P6	2.9 ± 0.1	32.65 ± 2.76	28.85 ± 3.38	19.09 ± 3.36	47.95 ± 4.99
P7	1.8 ± 0.1	21.38 ± 3.52	0.37 ± 0.02	0.25 ± NA	0.62 ± NA
P8	5.1 ± 0.5	35.5 ± 2.23	8.78 ± 0.99	10.81 ± 1.74	19.59 ± 2.11
P9	6.5 ± 0.5	33.24 ± 2.35	2.23 ± 0.19	0.28 ± 0.03	2.51 ± 0.2
P10	6	47.6	16.34	16.41	32.74
P11	135.2 ± 18.1	18.53 ± 1.97	9.91 ± 1.95	4.6 ± 1.99	14.51 ± 2.83

Table A2.4. Production rates of *n*-butyrate, *n*-caproate, and *n*-caprylate. Total production rate is sum of effluent and transfer rate. C8 to C6 production ratio is also reported. In Period 7 transfer rate for both *n*-caproate and *n*-caprylate was negative based on stripping data so set to zero. *n*-Caprylate effluent measurement only had one value. Therefore could not calculate C8 to C6 ratio. In Period 9 stripping was off. In this table, periods are abbreviated as P1 to P11 representing Periods 1 to 11. Mean ± standard error is reported.

	***n*-Butyrate Production Rates (g COD L^{-1} d^{-1})**			***n*-Caproate Production Rates (g COD L^{-1} d^{-1})**			***n*-Caprylate Production Rates (g COD L^{-1} d^{-1})**		
	Effluent	**Transfer**	**Total**	**Effluent**	**Transfer**	**Total**	**Effluent**	**Transfer**	**Total**
P1	0.63 ± 0.06	NA	0.63 ± 0.06	0.96 ± 0.04	3.25 ± 0.22	4.21 ± 0.23	0.44 ± 0.02	12.36 ± 1.16	12.8 ± 1.16
P2	0.49 ± 0.12	0.07 ± 0.03	0.56 ± 0.12	0.64 ± 0.1	1.19 ± 0.11	1.83 ± 0.15	0.41 ± 0.02	10.43 ± 1.03	10.84 ± 1.03
P3	0.31 ± 0.03	0.22 ± 0.11	0.53 ± 0.11	0.57 ± 0.05	4.15 ± 0.53	4.72 ± 0.54	0.25 ± 0.03	8.95 ± 1.69	9.2 ± 1.69
P4	1.24 ± 0.18	1.51 ± 0.21	2.75 ± 0.27	0.55 ± 0.05	9.38 ± 0.48	9.93 ± 0.48	0.1 ± 0.04	6.57 ± 0.64	6.67 ± 0.64
P5	2.46 ± 0.11	3.42 ± 0.45	5.88 ± 0.47	0.2 ± 0.02	7.03 ± 0.96	7.22 ± 0.96	NA	4.52 ± 0.98	4.52 ± 0.98
P6	2.67 ± 0.19	5.37 ± 0.55	8.04 ± 0.58	0.11 ± 0.01	7 ± 0.77	7.11 ± 0.77	NA	4.7 ± 0.8	4.7 ± 0.8
P7	2.44 ± 0.14	2.55 ± 0.79	4.99 ± 0.8	0.09 ± 0	0	0.09 ± 0	0.06	0	0.06
P8	3.94 ± 0.24	4.39 ± 0.22	8.33 ± 0.32	0.12 ± 0.01	1.94 ± 0.21	2.06 ± 0.21	0.06 ± NA	2.47 ± 0.39	2.54 ± 0.39
P9	7.98 ± 0.36	NA	7.98 ± 0.36	0.54 ± 0.03	NA	0.54 ± 0.03	0.07 ± 0.01	NA	0.07 ± 0.01
P10	4.38 ± 0.36	7.04 ± 1.99	11.42 ± 2.02	0.06 ± 0	3.86 ± 0.04	3.92 ± 0.04	NA	3.94 ± 1.23	3.94 ± 1.23
P11	1.23 ± 0.08	3.63 ± 0.44	4.86 ± 0.45	0.14 ± 0.04	2.46 ± 0.49	2.6 ± 0.49	0.08 ± 0	1.13 ± 0.52	1.21 ± 0.52

Table A2.5. Extraction efficiencies *per* period (mean ± s.e.) Extraction system was off for Period 9.

	Average Extraction Efficiency (%)		
Period	***n*-Butyrate**	***n*-Caproate**	***n*-Caprylate**
Period 1	NA	77.25 ± 6.71	96.6 ± 12.63
Period 2	11.75 ± 5.61	65.18 ± 7.93	96.23 ± 13.23
Period 3	41.6 ± 22.62	87.98 ± 15.07	97.27 ± 25.55
Period 4	54.86 ± 9.25	94.46 ± 6.6	98.5 ± 13.38
Period 5	58.13 ± 8.97	97.28 ± 18.56	100 ± 30.72
Period 6	66.77 ± 8.36	98.49 ± 15.23	100 ± 24.12
Period 7	51.02 ± 17.84	NA	NA
Period 8	52.67 ± 3.38	94.24 ± 13.91	97.53 ± 21.39
Period 9	NA	NA	NA
Period 10	61.65 ± 20.59	98.53 ± 1.46	100 ± 44.19
Period 11	74.63 ± 11.48	94.73 ± 26.15	93.71 ± 58.95

Table A2.6. Alpha diversity metrics for the bioreactor in this study (bottom and middle of bioreactor sampling locations), as well as the metrics for the four anaerobic digesters (AD) discussed in Chapter 3. 100 rarefactions at a depth of 6510 sequences *per* sample were performed and collated. Mean ± standard error is reported.

Bioreactor	**Bioreactor Position**	**Gini Coefficient**	**Observed OTUs**	**Shannon Diversity**
this study	Bottom	0.984±0.0	131±2	4.634±0.054
this study	Middle	0.987±0.0	129±3	3.973±0.09
AD 1	NA	0.989±0.0	366±6	5.366±0.057
AD 2	NA	0.988±0.0	373±5	5.491±0.056
AD 3	NA	0.988±0.0	368±7	5.505±0.046
AD 4	NA	0.989±0.0	365±8	5.375±0.07

Table A2.7. Gas Production Rate *per* period (for the main periods in the study only an average rate across all periods is reported; Period 1 is not included because data was not available). Mean ± s.e. is reported.

Period	Gas Production Rate (mL d^{-1})
Main (Periods 2 to 7)	0.38±0.01
Period 8	1.22±0.06
Period 9	1.19±0.03
Period 10	1.66±0.27
Period 11	2.74±0.04

Table A2.8. Gas composition in headspace of the bioreactor per period on a volumetric basis. Mean and standard deviation are reported. Measurements that were below detection limit (0.1%) are indicated by nd. The number of sampling days per period is noted by n. On each sampling day, duplicate or triplicate measurements were made of the bioreactor headspace. These measurements were averaged prior to finding the average per period.

Period	**Component**	**Composition % (mean ± s.d.)**
P1	Carbon dioxide	1.3 ± 0.8 (n=2)
	Hydrogen	0.7 ± 0.6 (n=2)
	Methane	16.5 ± 3.6 (n=2)
P2	Carbon dioxide	2.2 ± 0.2 (n=2)
	Hydrogen	0.2 ± 0.0 (n=2)
	Methane	21.7 ± 0.2 (n=2)
P3	Carbon dioxide	2.1 ± 0.1 (n=3)
	Hydrogen	0.3 ± 0.6 (n=3)
	Methane	21.3± 3.8 (n=3)
P4	Carbon dioxide	nd
	Hydrogen	nd
	Methane	17.6 (n=1)
P5	Carbon dioxide	1.3 ± 1.1 (n=3)
	Hydrogen	nd
	Methane	20.0 ± 1.4 (n=3)
P6	Carbon dioxide	2.0 (n=1)
	Hydrogen	nd
	Methane	32.2 (n=1)
P7	Carbon dioxide	2.1 ± 0.4 (n=2)
	Hydrogen	nd
	Methane	22.7 ± 1.3 (n=2)

Period	Component	Composition % (mean ± s.d.)
P8	Carbon dioxide	1.7 ± 0.7 (n=5)
	Hydrogen	0.1 ± 0.1 (n=5)
	Methane	19.8 ± 4.5 (n=5)
P9	Carbon dioxide	2.2 (n=1)
	Hydrogen	0.2 (n=1)
	Methane	47.2 (n=1)
P10	Carbon dioxide	2.1 (n=1)
	Hydrogen	nd
	Methane	56.9 (n=1)
P11	Carbon dioxide	1.9 ± 0.0 (n=4)
	Hydrogen	nd
	Methane	25.6 ± 2.8 (n=4)

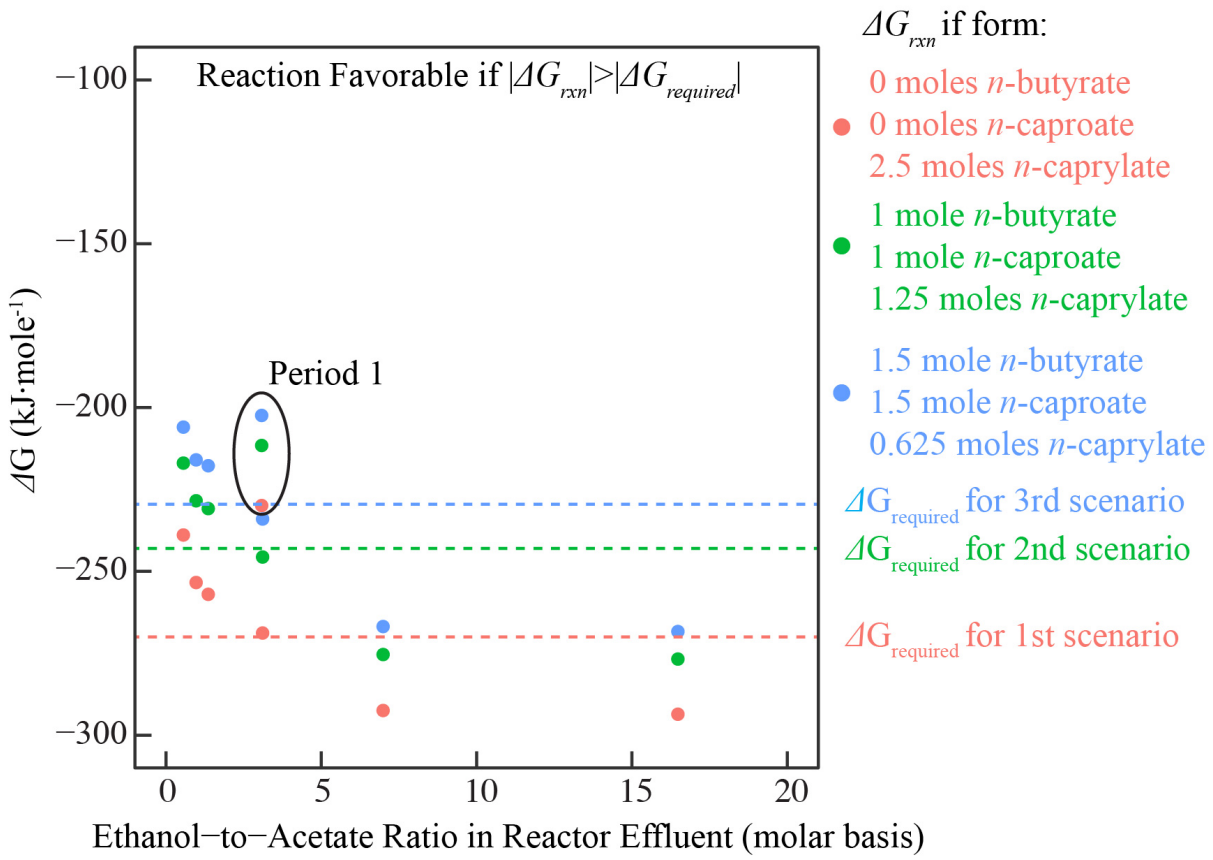

Figure A2.1. Gibbs free energy of the reaction *vs*. ethanol-to-acetate molar ratio measured in bioreactor effluent. Gibbs free energy of the reaction based on the ethanol and carboxylate concentrations measured in the bioreactor is plotted as ΔG of the reaction for the three different scenarios shown on the right of the figure. ΔG required is also plotted and was calculated based on the amount of ATP produced based on the stoichiometric model and assuming that -72 kJ is required *per* mole of ATP produced. Reaction was deemed thermodynamically favorable if the ΔG of the reaction was more negative than ΔG required. Period 1 data points are circled in the figure.

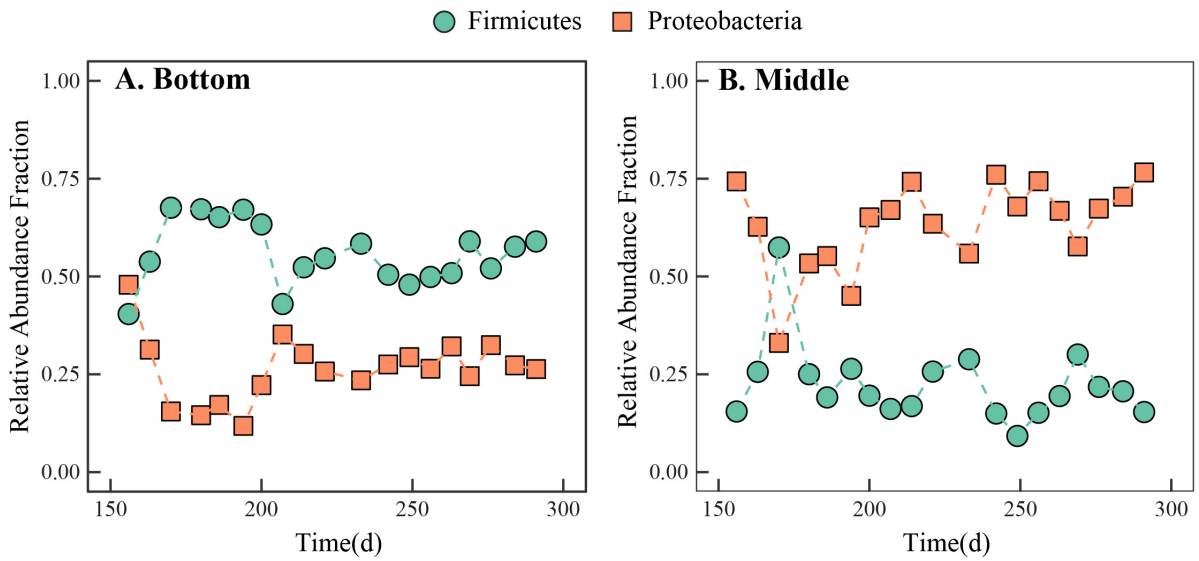

Figure A2.2. Relative abundance of the phyla Firmicutes and Proteobacteria in the samples collected from the bottom (**A**) and middle (**B**) sampling ports of the bioreactor.

APPENDIX 3

SUPPLEMENTAL INFORMATION FOR: VARIABILITY OF THE MICROBIOME INVOLVED IN *N*-CAPRYLATE PRODUCTION IN ANAEROBIC BIOREACTORS WITH IN-LINE PRODUCT EXTRACTION

Table A3.1. Conversion efficiency for *n*-caproate, *n*-caprylate, and the combined medium-chain carboxylic acids (MCCAs) per reactor and per period. Mean ± s.e. is reported.

		Conversion Efficiency (%)		
	Period	***n*-Caproate**	***n*-Caprylate**	**Total MCCA**
Reactor 1	Period 1	6±1	53±5	58±5
	Period 2	15±2	27±3	42±3
Reactor 2	Period 1	23±2	38±4	60±5
	Period 2	16±1	29±2	45±3
Reactor 3	Period 1	13±1	18±1	31±2
	Period 2	14±1	35±2	49±2

Table A3.2. MCC ratio (g COD of *n*-caprylate produced vs g COD of *n*-caproate produced) and specificity of carboxylates per period and per reactor. Mean ± s.e. is reported.

			Specificity (%)		
	Period	**MCC ratio**	***n*-Butyrate**	***n*-Caproate**	***n*-Caprylate**
Reactor 1	Period 1	9.4±0.6	1±0	10±2	89±7
	Period 2	1.9±0.2	5±1	33±4	62±7
Reactor 2	Period 1	1.6±0.1	3±1	37±3	60±5
	Period 2	1.7±0.1	4±0	35±2	61±4
Reactor 3	Period 1	1.4±0.1	9±1	39±2	52±4
	Period 2	2.4±0.1	2±0	28±1	69±2

Table A3.3. Effluent production rate, transfer rate (*via* pertraction) and total production rate for *n*-butyrate, *n*-caproate, and *n*-caprylate per reactor per period. Mean ± s.e. is reported. Periods 1 and 2 are abbreviated as P1 and P2. Reactors 1, 2, and 3 are abbreviated as R1, R2, and R3.

		***n*-Butyrate Production ($g\ COD\ L^{-1}\ d^{-1}$)**			***n*-Caproate Production ($g\ COD\ L^{-1}\ d^{-1}$)**			***n*-Caprylate Production ($g\ COD\ L^{-1}\ d^{-1}$)**		
		Effluent	**Transfer**	**Total**	**Effluent**	**Transfer**	**Total**	**Effluent**	**Transfer**	**Total**
R1	P1	0.04±0.00	0.01±0.01	0.05±0.01	0.06±0.00	0.35±0.06	0.41±0.06	0.06±0.00	3.77±0.23	3.83±0.23
	P2	0.10±0.01	0.08±0.01	0.18±0.02	0.10±0.01	1.00±0.12	1.11±0.12	0.03±0.00	2.04±0.18	2.07±0.18
R2	P1	0.07±0.01	0.07±0.02	0.14±0.02	0.09±0.01	1.61±0.11	1.70±0.11	0.03±0.00	2.77±0.20	2.80±0.20
	P2	0.04±0.00	0.10±0.01	0.14±0.01	0.05±0.00	1.15±0.04	1.20±0.05	0.02±0.00	2.07±0.11	2.09±0.11
R3	P1	0.08±0.01	0.15±0.02	0.23±0.02	0.05±0.01	0.97±0.05	1.03±0.05	0.02±0.00	1.37±0.08	1.39±0.08
	P2	0.09±0.01	NA	0.09±0.01	0.12±0.01	0.96±0.03	1.08±0.03	0.04±0.00	2.57±0.05	2.61±0.05

Table A3.4. Effluent ethanol and carboxylate concentrations per period. Mean and standard error reported.

		Ethanol (mM)	Acetate (mM)	*n*-Butyrate (mM)	*n*-Caproate (mM)	*n*-Caprylate (mM)
Reactor 1	Period 1	47.5±4.1	4.2±0.5	2.2±0.1	2.2±0.1	1.7±0.1
	Period 2	73.9±4.1	8.0±1.1	5.8±0.5	3.6±0.2	0.8±0.1
Reactor 2	Period 1	29.1±4.3	9.6±0.8	3.8±0.3	3.2±0.2	0.9±0.1
	Period 2	115.1±9.1	13.8±1.0	2.3±0.1	1.8±0.2	0.5±0.0
Reactor 3	Period 1	163.9±9.9	2.6±0.4	4.0±0.5	1.6±0.2	0.4±0.0
	Period 2	98.2±5.9	3.4±0.5	4.7±0.5	4.0±0.3	1.0±0.1

Table A3.5. Percent hydrogen (by volume) measured in the headspace of the three bioreactors. Mean and s.d. values are reported per period. Periods 1 and 2 are abbreviated as P1 and P2.

	Period	Hydrogen (%)
Reactor 1	P1	11.1±4.0
	P2	0.7±0.2
Reactor 2	P1	2.1±1.8
	P2	13.2±5.3
Reactor 3	P1	31.3±9.6
	P2	7.3±4.6

Table A3.6. Average total protein concentrations (mean ± s.d.) for the three bioreactors in Periods 1 and 2. Values reported are averages of three to four samples taken during period. For each time point, three separate dilutions (1x, 2x, and 5x) of the sample were prepared, measured, and averaged.

Reactor	**Period**	**Total Protein Concentration (μg mL^{-1})**
Reactor 1	Period 1	796.84±104.46 (n=3)
	Period 2	183.66±41.48 (n=4)
Reactor 2	Period 1	422.69±313.72 (n=4)
	Period 2	395.32±221.34 (n=4)
Reactor 3	Period 1	387.09±284.24 (n=4)
	Period 2	188.94±63.16 (n=4)

Table A3.7. Average Gini Coefficient, observed OTUs, and Shannon Diversity for the three bioreactors for Periods 1 and 2 of the study. Mean ± standard error is reported.

Reactor	Period	Gini Coefficient	Observed OTUs	Shannon Diversity
Reactor 1	Period 1	0.995±0.0	77±2	2.245±0.119
	Period 2	0.989±0.001	94±2	3.966±0.178
Reactor 2	Period 1	0.996±0.0	67±4	1.9±0.212
	Period 2	0.992±0.001	84±4	3.498±0.102
Reactor 3	Period 1	0.990±0.001	98±3	3.792±0.276
	Period 2	0.991±0.001	98±4	3.407±0.255

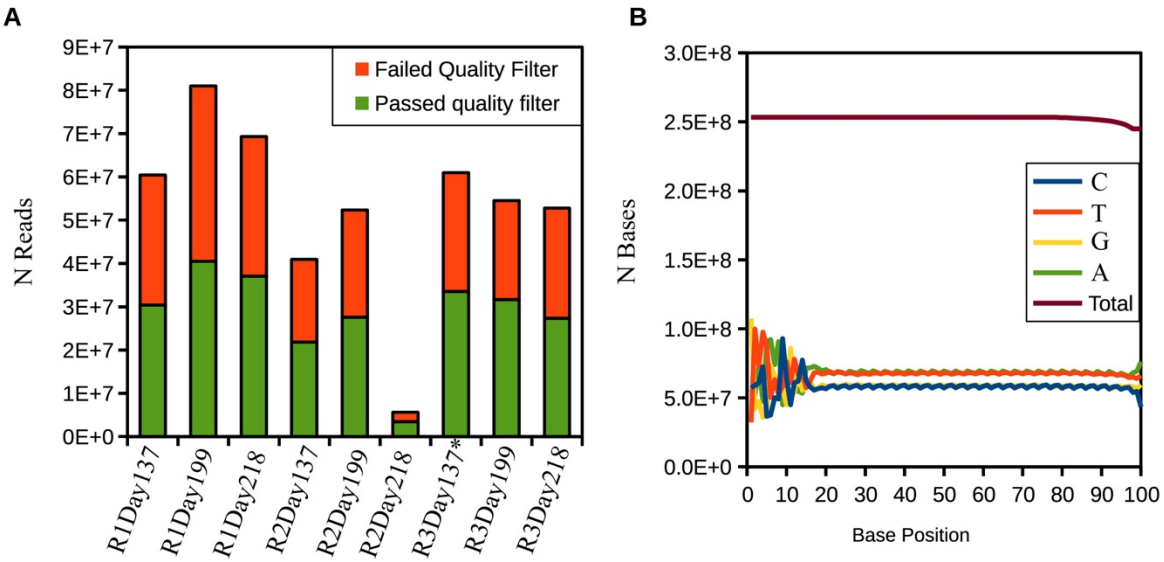

Figure A3.1. Number of reads per sample, and the proportion that passed the quality filtering step (**A**), and base pair frequencies at each base position throughout the raw reads, after quality filtering (**B**) for the shotgun metagenomics analysis. In panel **A**, R1 indicates Reactor 1 samples; R2 indicates R2 samples; and R3 indicates Reactor 3 samples.

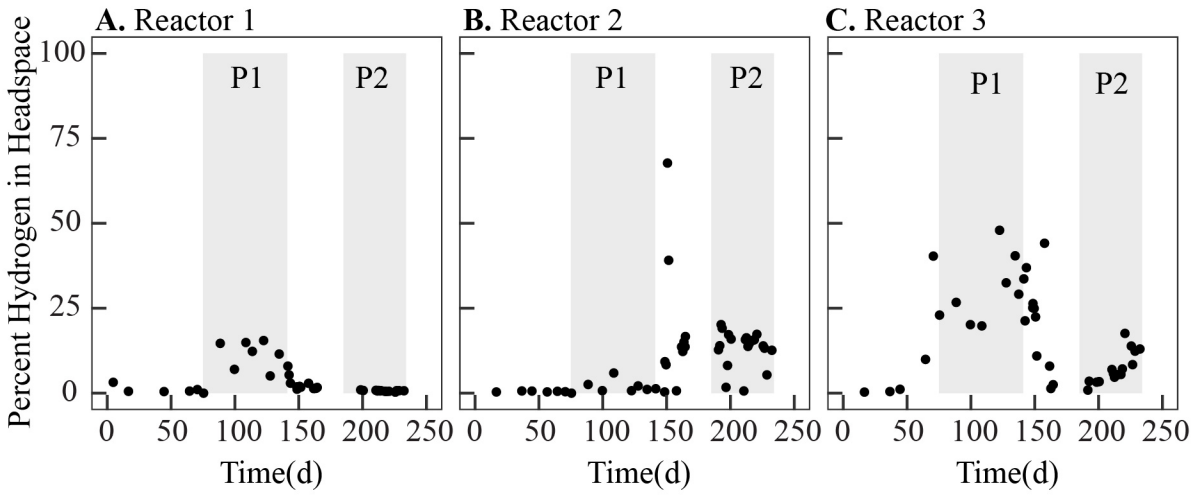

Figure A3.2. Percent hydrogen measured in headspace of bioreactors over the operating period for Reactor 1 (**A**), Reactor 2 (**B**), and Reactor 3 (**C**). Periods 1 and 2 (P1 and P2) are shaded in the Figure.

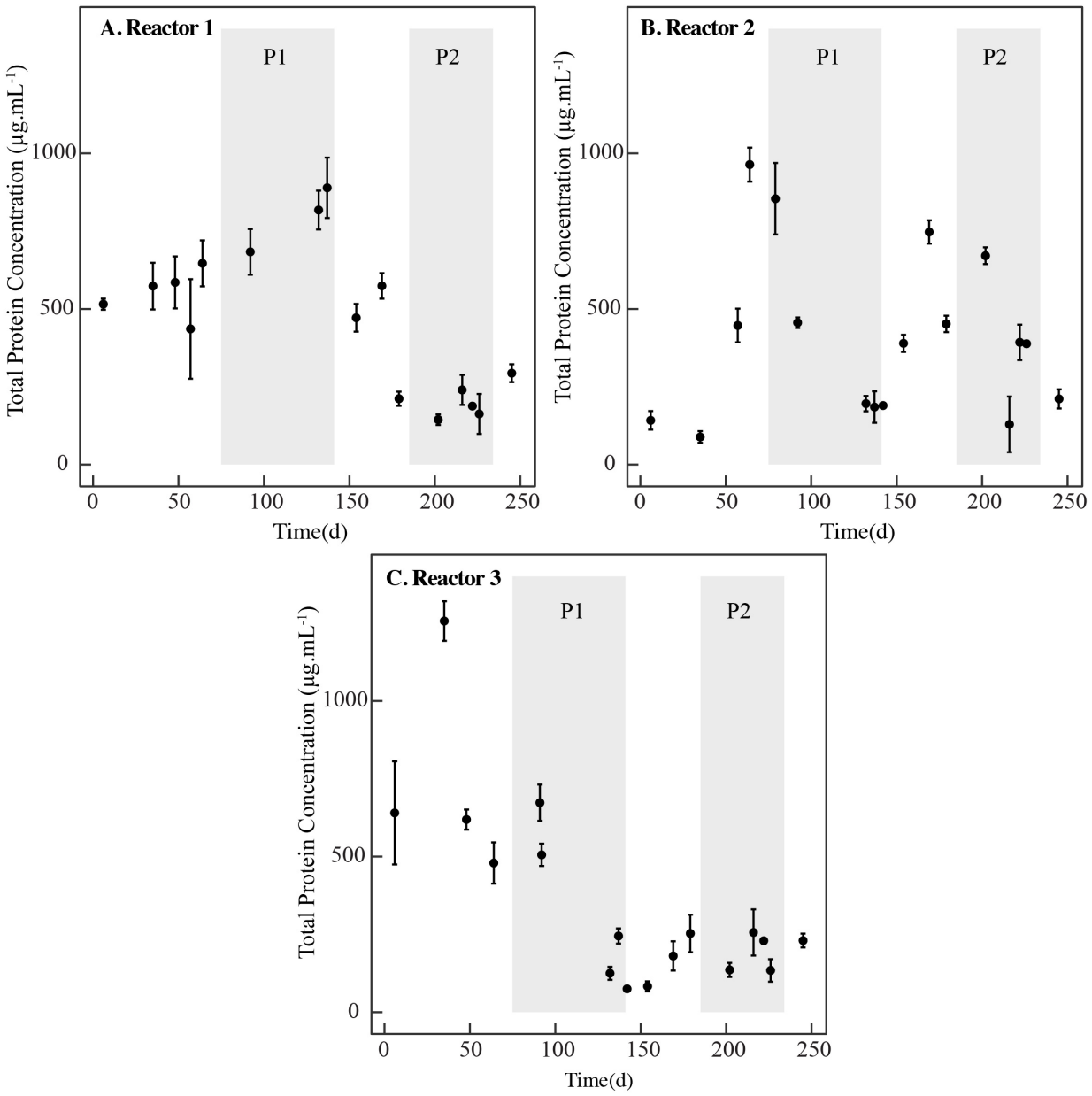

Figure A3.3. Average total protein concentrations in the bioreactors over the operating periods for Reactor 1 (**A**), Reactor 2 (**B**), and Reactor 3 (**C**). Periods 1 and 2 are shaded in the Figure. Error bars represent standard deviation (n=3).

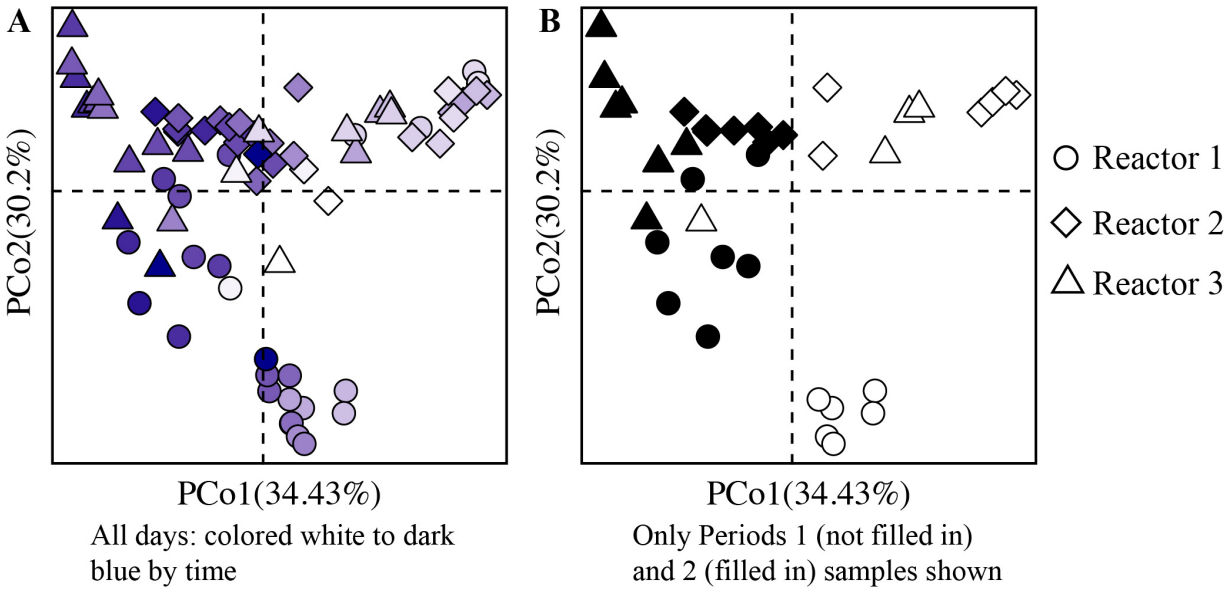

Figure A3.4. Principal coordinates analysis based on weighted Unifrac distance of biomass samples from Reactors 1, 2, and 3 during the operating period. The first two principal coordinate axes are shown. All biomass samples are shown in panel **A** (from startup to last sampling day) and the samples are colored white to dark blue with increasing time. In panel **B**, the same principal coordinates analysis plot is shown but only the samples from Period 1 (shapes not filled in) and Period 2 (shapes filled in) are displayed on the plot.

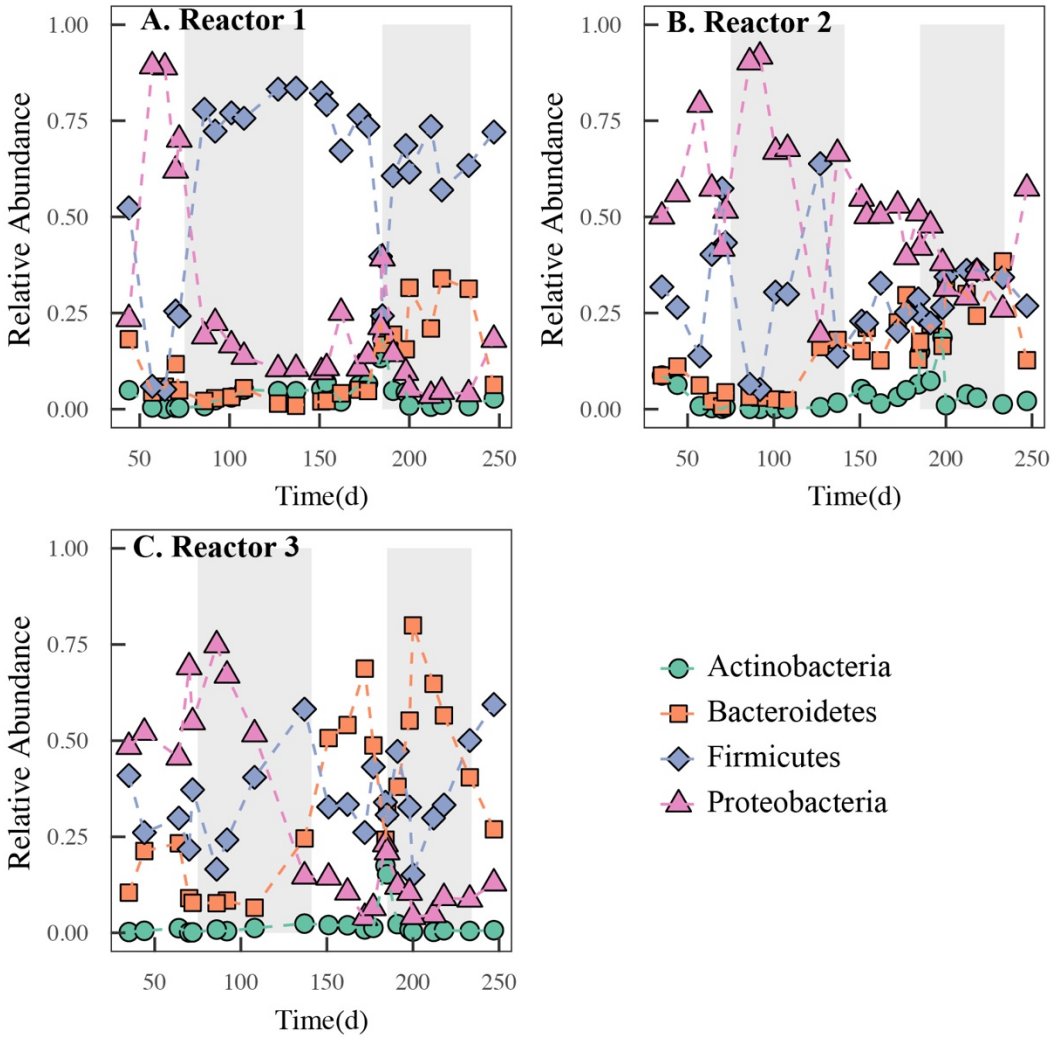

Figure A3.5. Relative abundance fraction during the operating period of the four phyla which reach over 5 percent relative abundance (over 0.05 fraction) in any one bioreactor sample. Plots are for **A**) Reactor 1, **B**) Reactor 2; and **C**) Reactor 3 microbiomes. Shading represents Periods 1 and Periods 2.

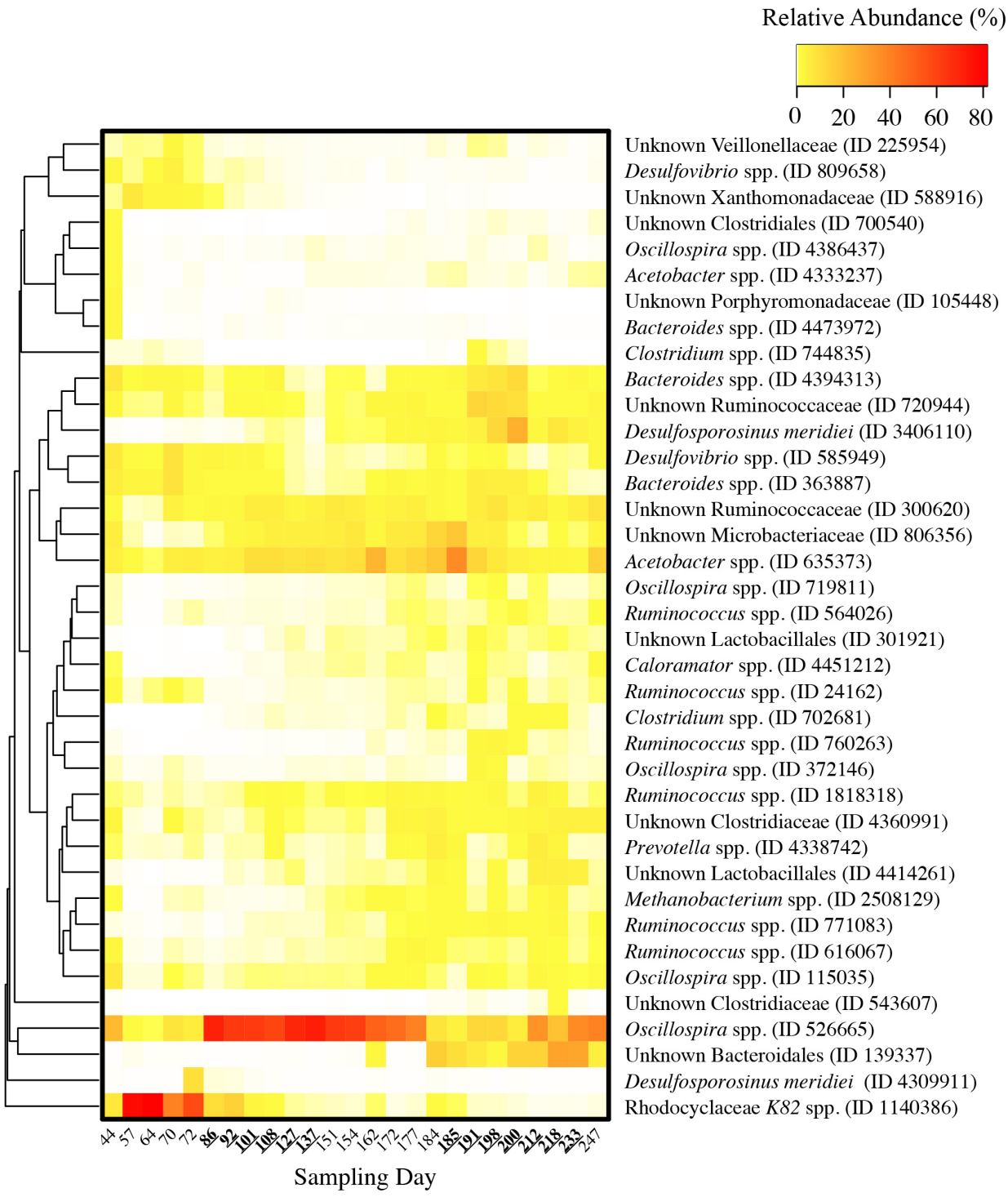

Figure A3.6. Heat map of OTUs that reached higher than one percent relative abundance in Reactor 1 over the main operating period. Lowest level taxonomy is reported. OTUs are clustered hierarchically (average linkage) based on the Bray-Curtis dissimilarity index.

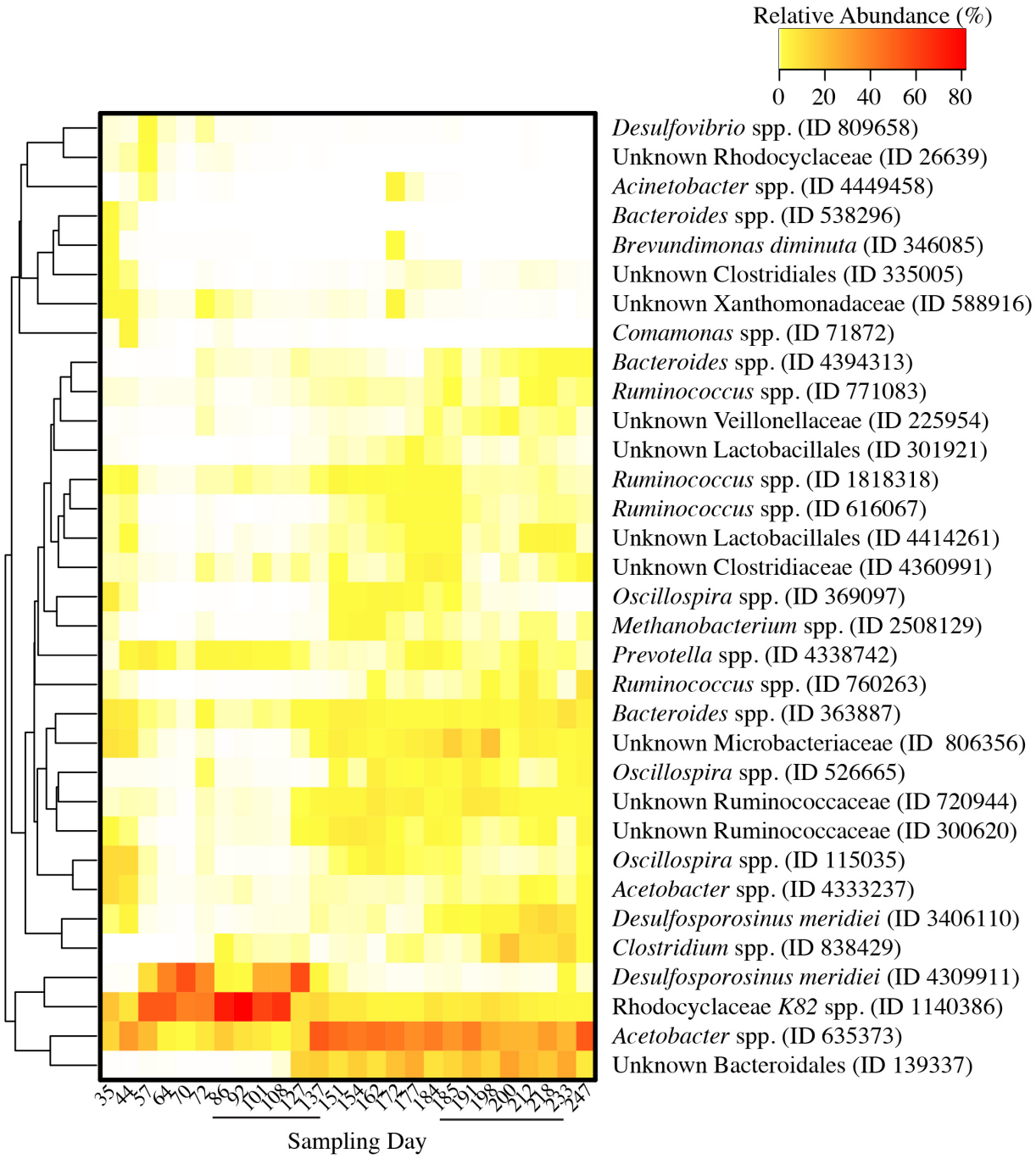

Figure A3.7. Heat map of OTUs that reached higher than one percent relative abundance in Reactor 2. Lowest level taxonomy is reported. OTUs are clustered hierarchically (average linkage) based on the Bray-Curtis dissimilarity index.

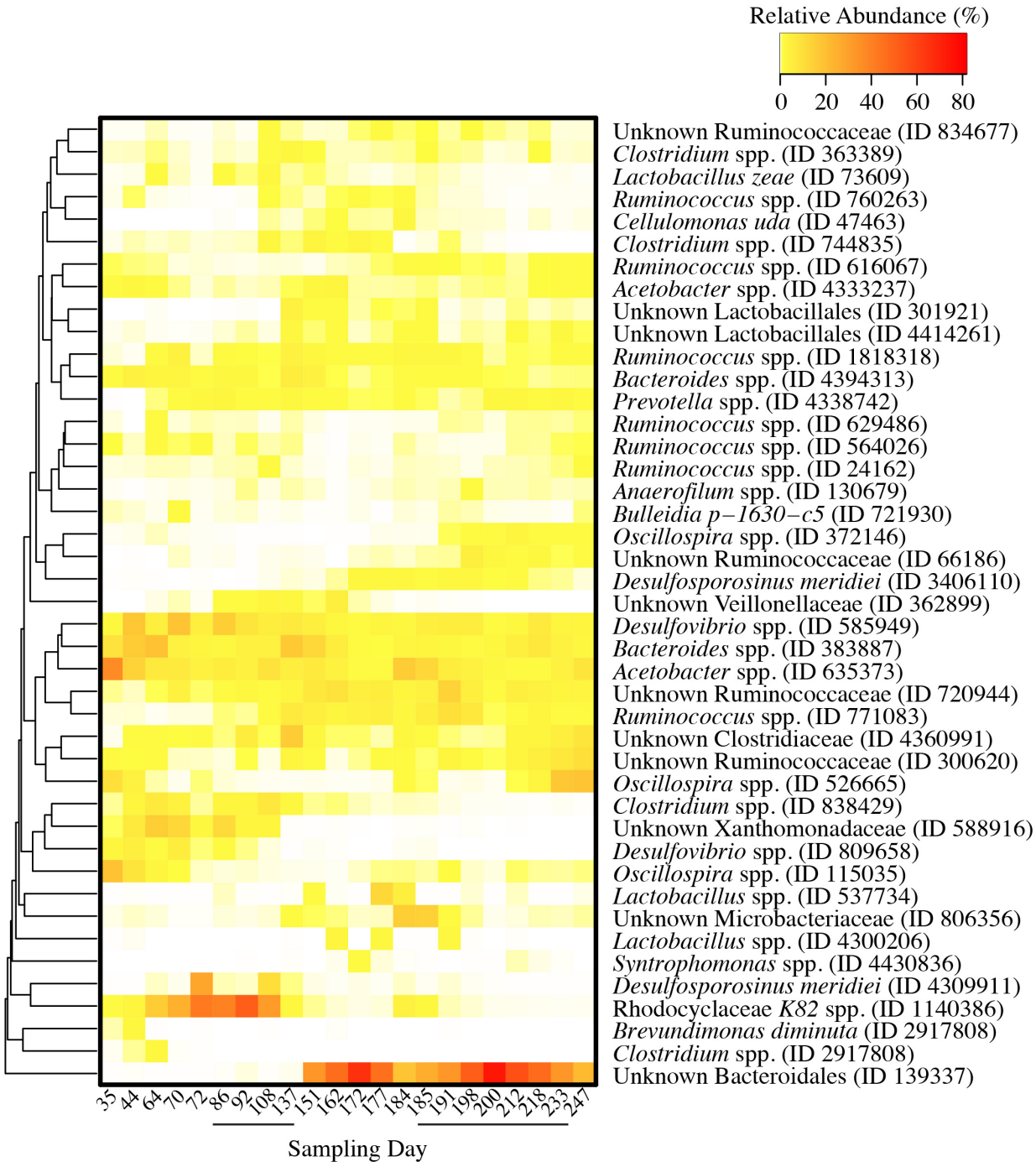

Figure A3.8. Heat map of OTUs that reached higher than one percent relative abundance in Reactor 3. Lowest level taxonomy is reported. OTUs are clustered hierarchically (average linkage) based on the Bray-Curtis dissimilarity index.

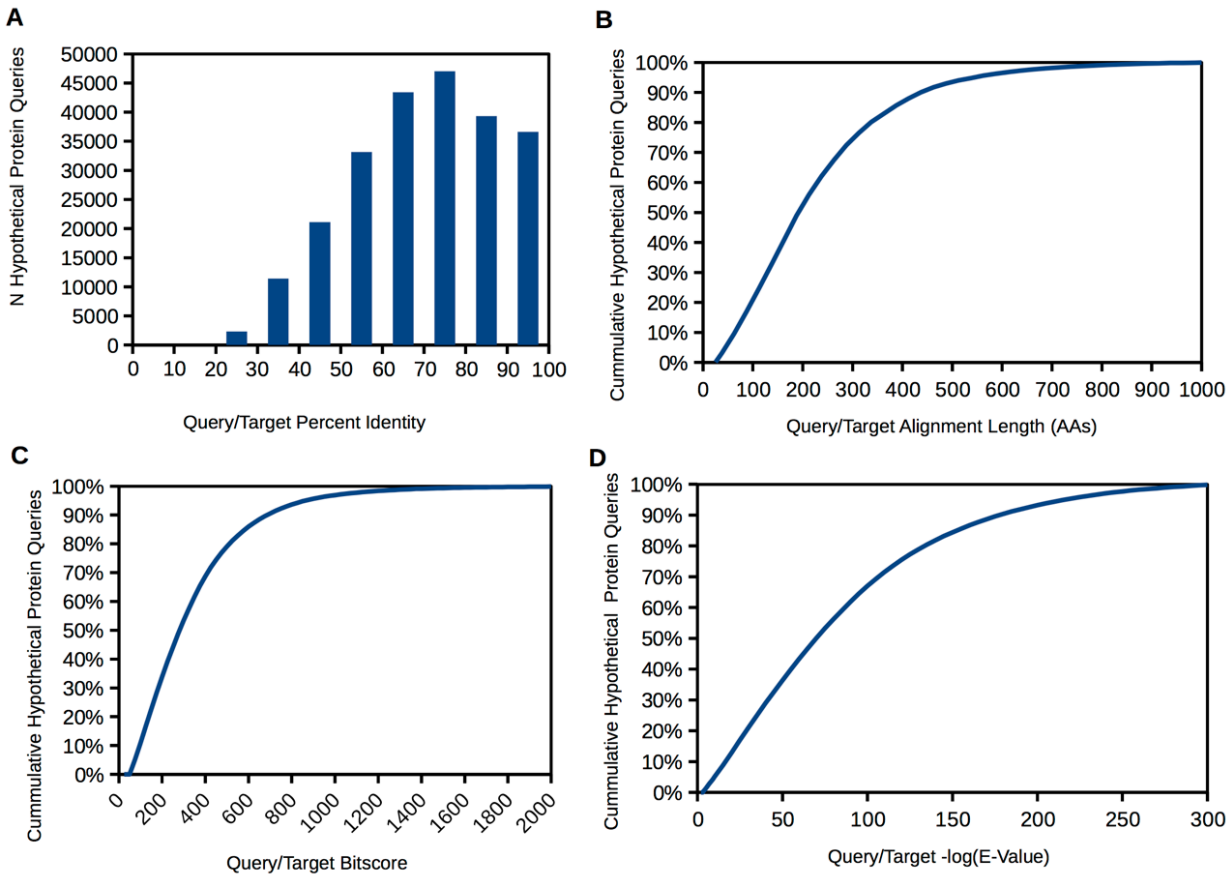

Figure A3.9. Hypothetical protein-coding genes from the metagenome translated to protein sequences that matched well with reference protein sequences on NCBI-nr, including an average percent identity of 72% (**A**), a median query/target alignment length of 204 amino acids (**B**), a median query/target bitscore of 281 (**C**), and a median e-value of 10^{-70} (**D**); stats for e-values performed on a logarithmic scale).

APPENDIX 4

PROTOCOLS

4.1. Sequencing of Environmental Samples Using Illumina MiSeq: Preparation and Data Processing

Originally compiled by Elliot Friedman and edited by Catherine Spirito and Mytien Nguyen in the Angenent Lab at Cornell University

Materials:

PowerSoil® DNA Isolation Kit

515f Forward primer

806r Barcoded reverse primer

5PRIME HotMaster Mix

Molecular grade H_2O

96 well PCR plates

Mag-Bind RxnPure Plus (M1366-01 from Omega)

96 well Mag-Bind PCR Cleanup Plate

70% molecular grade ethanol (from at least 200 proof (VWR #459844)– no methanol allowed)

Magnetic separation stand

Elution Buffer (Solution C6 from PowerSoil® DNA Isolation Kit)

PicoGreen Reagent

Tris EDTA solution

Polypropylene reservoirs

Various single- and multi-channel pipettes

0.1-10 µL filter pipette tips

200 µL filter pipette tips

Eppendorf pipette tips

Procedure:

1. Extract genomic DNA from samples using the PowerSoil DNA Isolation Kit. Procedure: http://www.mobio.com/images/custom/file/protocol/12888.pdf
2. Run duplicate 50 µL polymerase chain reactions (PCRs) for each sample by combining (for each reaction): 0.5 µL of 10 µM 515f forward primer, 0.5 µL of 10 µM 806r barcoded reverse primer, 20 µL of 5PRIME HotMaster Mix, 28 µL of molecular grade H_2O, and 1 µL of DNA template (extraction product). The PCR sequence is titled 16SEMP25 on the thermocycler. The thermocycler program is as follows:
    a. 94°C for 3 min
    b. 94°C for 45 s
    c. 50°C for 60 s
    d. 72°C for 90 s
    e. Repeat steps b-d 24 times (for a total of 25 PCR cycles)
    f. 72°C for 10 min
    g. 4°C hold

    Note: for low DNA samples, perform PCR with 30-35 cycles, but only go to 35 cycles when you have no other choice. It is problematic and all blanks must be negative. With 35 cycles it is likely that amplicons appear from just the polymerase or other chemicals and clean water. A constant check is needed. A different batch of enzymes can change the outcomes for the negative.

    Note: details on the forward primer and reverse barcoded primer sequences used can be found at http://www.earthmicrobiome.org/emp-standard-protocols/16s/
3. For each sample, combine the PCR product from duplicate reactions. *At this step if not running gel immediately, aliquot minimum of 5 µL to run gel later (into individual PCR tubes).
4. Run gels to confirm PCR product.

5. Purify PCR product with Mag-Bind RxnPure Plus.
    a. Shake Mag-Bind Solution to resuspend any settled particles. Add 1.8X PCR product (1.8*80 µL) of Mag-Bind to each well. [In DNA biosafety cabinet, rest of protocol outside hood in clean bench area]
    b. Transfer 80 µL of pooled PCR product to 96-well Mag-Bind PCR Cleanup Plate
    c. Mix each well by pipetting up and down 5-10 times (or vortexing for 30 s)
    d. Place the plate onto a magnetic separation stand to magnetize the Mag-Bind particles. Solution will be clear when beads have completely migrated toward the magnets [at least 5 min]. Leave plate on stand.
    e. Remove and discard the clear supernatant (don't disturb magnetic beads, should remove ~200 µL).
    f. Add 200 µL of 70% ethanol (made with 200 proof ethanol) to each well (don't need to resuspend particles) and incubate at room temperature for 1 min.
    g. Remove and discard the clear supernatant (don't disturb magnetic beads, remove ~200 µL).
    h. Repeat steps f and g.
    i. Allow the plate to dry on the magnetic for at least 10-15 min. Remove any liquid residue from the wells by pipetting [Make sure it is dry before proceeding; don't move plate at this step, the beads can be easily disturbed].
        i. Go back after 1 min to remove residual with 10 µL pipette
        ii. You can use breath-easy plate seal to cover wells during this step but be very careful when removing (beads may stick to seal)
    j. Remove from the separation stand **(carefully!)** and add 30-40 µL of Elution Buffer (10 mM Tris pH 8.5, TE Buffer, 0.1 mM EDTA, DI H2O, or Solution C6 from PowerSoil® DNA Isolation Kit) to each well to elute DNA from the magnetic particles. [Make sure the volume is above the magnetic beads].
        i. Make sure volume is above ring
    k. Mix each well by pipetting up and down 20 times (or vortexing for 30 s), then incubate at room temperature for greater than 5 min.
    l. Place the plate onto the magnetic seperation stand to magnetize the Mag-Bind particles. Wait at least 5 min.
    m. Transfer the cleared supernatant containing purified DNA to a new 96-well PCR plate. [Use 10 µL pipette, if you draw beads, put solution back and let settle, then retry; you should get ~30-35 µL out]
    n. Seal the PCR plate with aluminum foil and store at -20^{o}C.
6. Run gels to confirm bands are expected size.
7. Quantify DNA using PicoGreen.
    a. Take out PicoGreen to thaw.
    b. Make linear DNA standards from lamda-DNA stock.
    c. You will need 99 µL of TE and 100 µL green mix per sample. To calculate the amount of green mix you need (# of samples ±1) x 100 = volume (µL) of green mix. To make green mix, make a 1:200 dilution of PicoGreen in TE. (For example, to make 1mL of green mix you would dilute 5µL PicoGreen in 995 µL sterile TE). Mix, and keep in the dark until ready to use.

d. Using a black 96-well plate, first add 99uL of TE to each well that a sample or standard will go in.
e. Add 1μL of sample or standard per well. Remember to have one 'blank' containing only TE and green mixture (no DNA).
f. Add 100uL of green mixture to each well. Pipette up and down to mix well.
g. Read fluorescence at Excitation 485/20; Emission 528/20; Sensitivity 50.
h. Calculate DNA concentrations in samples using standard curve.

8. Pool 100 ng DNA from each sample into a single tube. Aliquot the pooled samples into 50 μL volumes (only one 50 μL volume needs to be brought to the sequencing center, you can store the rest as backup)

   Notes (check with your own sequencing center):

   - The minimum volume accepted by the sequencing center is 15 μL
   - For low DNA concentrations, pool 20 ng DNA from each sample
   - Run gel on pooled sample to ensure correct size – if not, perform another round of PCR-cleanup

9. Sequence at the Cornell University Biotechnology Resource Center using the Illumina MiSeq (2x250bp; paired end). Specifically for Cornell:
   a. Create your own account on the BRC website by going to http://www.biotech.cornell.edu/brc/genomics-facility >> Sample Submission.
   b. To complete an order, first create a new project and experiment, then click Place Order.
      i. After you input your account details, you'll be asked about your order information. Select the following: Illumina MiSeq, 2x250bp paired-end, custom library submission, barcode primer length = 12
      ii. On the next page, note that the sample is a 16S rRNA (Rob Knight) primer amplicon in the comments box, then enter your sample name, pool = blank, Number of lanes = 1, volume of submission, and concentration of DNA.
   c. To get accurate DNA concentration prior to submission: use Qubit machine in Rm 152 of Biotech building
   d. Drop sample off at the BRC Genomics Diversity Facility (140 Biotech Bldg (*NOT* 147 Biotech)
10. A couple of days after, you'll receive a Fragment Analyzer QC data of your sample. Your target DNA should be around 380's bp. The main peaks to check for are significant DNA fragments between 130bp – 500bp that's not your target DNA. If there is, talk to a lab technician from the center (any technician listed under Genomics Diversity Facility in the BRC staff directory). You may need to repeat a PCR cleanup.
11. You will receive three files from the sequencing center: forward reads, reverse reads, and the 12 base barcode read. Join the forward and reverse reads using the joined_paired_ends.py script in QIIME and then upload to QIITA (qiita.microbio.me) for further processing.

4.2 16S and ITS Illumina MiSeq Analysis using QIITA

Protocol written by Mytien Nyugen of the Angenent Lab.

http://qiita.ucsd.edu or http://qiita.microbio.me

- Create an account by clicking “Sign Up”

BEFORE YOU BEGIN

Definitions

Sample template: a tab-delimited txt file containing information about your samples, including environmental and other important information about them. If you collected 100 samples, you will need 100 rows in your sample template describing each of them, including blanks. A template of this file is in the Angenent Lab Google drive.

Prep template: a tab-delimited txt file containing basic information about wet lab work on all or a subset of the samples. In your 100 samples, if you prepare 95 of them for 16S and 50 of them for ITS, you will need 2 prep templates: one with 95 rows describing the preparation for 16S, and another one with 50 to describing the ITS. A template of this file is in the Angenent Lab Google drive. A list of required columns and data for both the sample and prep templates are also detailed on the Qiita website (http://qiita.microbio.me/static/doc/html/tutorials/prepare-templates.html).

Pre-processing: includes the first steps in sequencing analysis, including demultiplexing and quality filtering. This step outputs quality filtered joined reads.

Processing: includes clustering, OTU picking, and assigning taxonomies, and outputs a biom OTU table.

About studies in Qiita

Studies can contain one set of samples or multiple sets of raw data, each of which can have a different preparation. The number of sample template and prep template you will need depends on the number of datasets you have. For example, imagine a study with 100 samples in which:

- Scenario 1: all of the samples were prepped for 16S and sequenced in a single MiSeq run
- Scenario 2: all of the samples were prepped for 16S and sequenced in two MiSeq runs
- Scenario 3: all of the samples were prepped for 16S, and 50 were also prepped for ITS. All 16S and ITS samples were sequenced in a single MiSeq run
- Scenario 4: all of the samples were prepped for 16S and sequenced in a single MiSeq run, and samples are from two separate studies (50 from study 1, 50 from study 2)

To represent this project in Qiita, you will need to create a single study with a single sample template that contains all 100 of the samples. Separately, you will need to create 4 prep templates that describe the preparations for the corresponding samples. All raw data uploaded will need to correspond to a specific prep template. For instance, the data sets described above would require the following data and template information:

Scenario 1:

- 1 prep template describing the MiSeq run where the 100 samples are represented

- The 3 fastq raw data files without demultiplexing from the sequencing center (*i.e.*, the forward-R1, reverse-R2, and barcodes-I1)

Scenario 2:

- 1 prep template describing the two MiSeq runs (use the run_prefix column to differentiate between the two MiSeq runs, corresponding to the file names) where the 100 samples are represented
- The 6 fastq raw data files without demultiplexing from the sequencing center (*i.e.*, 2 files for forward-R1, 2 files for reverse-R2, and 2 files for barcodes-I1)

Scenario 3:

- 2 prep templates, one describing the 16S preparations and the other describing the ITS preparations
- The 3 fastq raw data files without demultiplexing from the sequencing center (*i.e.*, the forward-R1, reverse-R2, and barcodes-I1)

Scenario 4:

- 2 prep templates, one for each study describing the MiSeq run where the study's 50 samples are represented
- The 3 fastq raw data files without demultiplexing from the sequencing center (*i.e.*, the forward-R1, reverse-R2, and barcodes-I1)

Study status

Sandbox: When a study is in this status, all the required metadata columns must be present in the metadata files (sample and prep), but the values don't have to be filled in or finalized yet. The purpose of this status is so that users can quickly upload their sequence files and some (possibly incomplete) metadata to have a preliminary look at their data.

Private: Moving from sandbox to private status requires the user to correct and finalize their metadata. On the each study overview page, there is a button that the user can use to request approval. Approval must be provided by a Qiita admin, who will validate and finalize the metadata. After a study moves from sandbox to private status, very little can be changed about the study without reverting the study to sandbox.

Public: Once a study is made administrator-approved and becomes private, the user can choose when to make it public. Making a study public means that it will be available to anyone with a Qiita user account (e.g., for data downloads and meta-analyses).

PROTOCOL

Creating a study

1. To create a study, click on the "Study" menu and then on "Create Study". This will take you to a new page that will gather some basic information to create your study.
2. Enter a unique "Study Title", and select the appropriate principal investigator(s) from the list. If the PI is not in this list, you can choose to add a new one. Note: Lars is on the list—in fact, he is on it twice—just pick one.
3. (Optional) select the environmental package appropriate to your study. Different packages will request different specific information about your samples.
4. Select the kind of time-series you have. The main options are:
    - No time-series: the samples do not represent a time-series.
    - Single intervention: the study has only one intervention, the classic before/after design. This can be also selected if you are only following individuals/environments over time without an actual intervention.
    - Multiple interventions: the study includes multiple interventions, such as 2-3 antibiotic (ABX) interventions.
    - Combo: the samples are a combination of those having single and multiple interventions.

    Additionally, there is a distinction between real, pseudo or mixed interventions:

    - Real: the study follows the same individuals over time, so there are multiple samples from the same individuals.
    - Pseudo: the study has time information from diverse individuals; for example, it includes samples from individuals from 3 to 60 years of age but has only one sample per individual.
    - Mixed: the study is a combination of real and pseudo.

Adding sample template

When you click on a study, you'll be taken to the study description page. Here, you will be able to edit the study info, upload files, and manage all other aspects of your study.

5. To upload your sample template, prep template, and sequence files, click on the "Upload" button. Drag-and-drop files into the grey area or simply click on "select from your computer" to select the fastq, fastq.gz or txt files you want to upload.
   Note: uploads can be paused at any time and restarted again, as long as you do not refresh or navigate away from the page, or log out of the system from another page.
6. Once your file(s) have been uploaded, from the upload tool, click on "Go to study description" and click on the "Sample template" tab. Select your sample template from the dropdown menu, then click "Process sample template". If it is processed successfully and is EBI-compliant, a green message will appear; if processing is successful, but is not EBI-compliant, a yellow message will appear; if processing is unsuccessful, a red message describing the errors will appear. If the last two, fix the described issues, re-upload your sample template file, and then re-attempt processing.
7. You can download the processed sample template under the "Sample template" tab once it has been successfully processed.

Linking raw data

(Video: 01 Adding Raw Files.MOV)

8. Under Data Type, click on '16S' (or 'ITS', depending on your prep), and click on the PREP link that appears. Check your Prep information. In 'Select type', select FASTQ-none. You will see your uploaded R1, R2, and I1 files. Drag and drop the files into the correct column designation—this is done correctly when all column backgrounds are green.
9. To add the files, click "Add files". This step will take ~30 min.
10. If you have more than one prepping templates, repeat this for the rest of your prepping templates. Each prepping file will have their own subtab with a unique ID, *i.e.*) 16S (ID: ###) or ITS (ID: ###)
11. Once you've successfully added your prepping template(s), you can download your prepping template, a summary of your prepping template, and a QIIME-compatible mapping file, which is a combination of your sample template and prepping template.

Preprocessing and processing your raw data

(Video: 02 Preprocessing.MOV and 03 Processing.MOV)

12. Once you have linked your raw data and your prep template has been processed, you can then proceed to preprocessing your data for each preparation by going back to your '16S' – PREP page under Data Types. Under file network, you should have a button named 'dflt_name-FASTQ'. Click on this button and click 'Process'. Here, you will define the preprocessing steps. Click on the bubble 'dflt_name-FASTQ' and scroll down. Select 'Split libraries FASTQ' for Choose command, dflt_name for input data, and the appropriate barcode for Choose parameter set. For the latter, if you have 16S 515F/806R Golay sequences, select the *Golay 12bp, reverse complement mapping file barcodes* option, with the following parameters:

    max_barcode_errors: 1.5

    sequence_max_n: 0

    max_bad_run_length: 3

    rev_comp: **False**

    phred_quality_threshold: 3

    rev_comp_barcode: **False**

    rev_comp_mapping_barcodes: **True**

    min_per_read_length_fraction: 0.75

    barcode_type: golay_12

13. Click the Add command, then Run.
14. Pre-processing may take a couple of hours to complete, and you may logout and do whatever in the meantime. The status is reported within your dflt_name – FASTQ file status page. Once completed, there'll be a second button under Files network, called dflt_name – Demultiplexed.
15. Click on the dflt_name – Demultiplexed button to see your preprocessed files. You will have 4 new files:

    *_seqs.fna: demultiplexed sequences in fasta format

    *_seqs.fastq: demultiplexed sequences in fastq format

    *_seqs.demux: demultiplexed sequences in an HDF5 format

*_split_library_log: the classic QIIME split libraries log that summarizes the demultiplexing process

16. If you are happy with the results, click on "Process" above your files to process your data. Under file network, click on the dflt_name – Demultiplexed button. Here, you will define the processing steps. Click on the bubble 'dflt_name-Demultiplexed' and scroll down. Select 'Pick closed-reference OTUs' for Choose command, dflt_name for input data, and 'default - serial' or 'default - parallel' for Choose parameter set. Check the processing parameters that appears (reference 1 = Greengenes 13_8), then click Add Command, then Run. Processing will take around 24 hours or more (depends on how complex your data is), so you should let it run overnight. You do not need to be logged in.
17. Once the processing is completed, a new button will be created under the Files network. You will have 3 new files:

    Biom – an OTU biom file of your data, in HD5 format (compatible with QIIME 1.9.0 and later)

    Directory – (not sure what this is)

    Log – a log file that summarizes the clustering, OTU picking, and taxonomic assignment processes ran in QIIME

18. If you have other prep templates, repeat steps 11-16 for each of them.
19. Complete your microbiome analysis in your local QIIME with the biom OTU table(s) and QIIME mapping file(s) from step 9

Analyzing your data with Qiita

20. Once your sequences have completed processing, you can conduct preliminary microbiome analysis on Qiita. Go to Study > View Studies and select the study(ies) you'd like to add to the analysis. Click Add to Analysis next to each study. Once you've successfully selected your samples, the Selected link on the top right corner should be green.
21. Click Analysis > Create from Selected samples. A list of your selected study will be visible. Here, you can remove samples and/or studies (samples list can be expanded by clicking Show/Hide Samples) you want to exclude from the analysis. Once you have your samples list, click Create Analysis.
22. Specify your rarefaction depth (you can leave it blank for no rarefying), and select any/all of the analysis types. Click Start Processing. Some analyses may take a while, and you can navigate out of the page while it is processing.

23. All analyses are listed under Analysis > See Previous Analyses.

REFERENCES

Abubackar, H.N., Bengelsdorf, F.R., Dürre, P., Veiga, M.C. and Kennes, C. (2016) Improved operating strategy for continuous fermentation of carbon monoxide to fuel-ethanol by clostridia. Applied Energy 169, 210-217.

Agler, M., Spirito, C., Usack, J., Werner, J. and Angenent, L. (2014) Development of a highly specific and productive process for *n*-caproic acid production: applying lessons from methanogenic microbiomes. Water Science & Technology 69(1), 62-68.

Agler, M.T. (2011) Improving performance and product specificity of the carboxylate platform for production of biochemicals, Cornell University, Ithaca, NY.

Agler, M.T., Spirito, C.M., Usack, J.G., Werner, J.J. and Angenent, L.T. (2012a) Chain elongation with reactor microbiomes: upgrading dilute ethanol to medium-chain carboxylates. Energy & Environmental Science 5(8), 8189-8192.

Agler, M.T., Werner, J.J., Iten, L.B., Dekker, A., Cotta, M.A., Dien, B.S. and Angenent, L.T. (2012b) Shaping reactor microbiomes to produce the fuel precursor *n*-butyrate from pretreated cellulosic hydrolysates. Environmental Science & Technology 46(18), 10229-10238.

Agler, M.T., Wrenn, B.A., Zinder, S.H. and Angenent, L.T. (2011) Waste to bioproduct conversion with undefined mixed cultures: the carboxylate platform. Trends in Biotechnology 29(2), 70-78.

Akuzawa, M., Hori, T., Haruta, S., Ueno, Y., Ishii, M. and Igarashi, Y. (2011) Distinctive responses of metabolically active microbiota to acidification in a thermophilic anaerobic digester. Microbial Ecology 61(3), 595-605.

Allison, S.D. and Martiny, J.B.H. (2008) Resistance, resilience, and redundancy in microbial communities. Proceedings of the National Academy of Sciences of the United States of America 105, 11512-11519.

Andel, J.G., Zoutberg, G.R., Crabbendam, P.M. and Breure, A.M. (1985) Glucose fermentation by *Clostridium butyricum* grown under a self generated gas atmosphere in chemostat culture. Applied Microbiology and Biotechnology 23(1), 21-26.

Andersen, S.J., De Groof, V., Khor, W.C., Roume, H., Props, R., Coma, M. and Rabaey, K. (2017) A *Clostridium group IV* species dominates and suppresses a mixed culture fermentation by tolerance to medium chain fatty acids products. Frontiers in Bioengineering and Biotechnology 5, 8.

Angenent, L.T., Karim, K., Al-Dahhan, M.H., Wrenn, B.A. and Domíguez-Espinosa, R. (2004) Production of bioenergy and biochemicals from industrial and agricultural wastewater. Trends in Biotechnology 22(9), 477-485.

Angenent, L.T., Mau, M., George, U., Zahn, J.A. and Raskin, L. (2008) Effect of the presence of the antimicrobial tylosin in swine waste on anaerobic treatment. Water Research 42(10-11), 2377-2384.

Angenent, L.T., Richter, H., Buckel, W., Spirito, C.M., Steinbusch, K.J.J., Plugge, C.M., Strik, D.P.B.T.B., Grootscholten, T.I.M., Buisman, C.J.N. and Hamelers, H.V.M. (2016) Chain elongation with reactor microbiomes: Open-culture biotechnology to produce biochemicals. Environmental Science & Technology 50(6), 2796-2810.

Antonopoulos, D.A., Huse, S.M., Morrison, H.G., Schmidt, T.M., Sogin, M.L. and Young, V.B.

(2009) Reproducible community dynamics of the gastrointestinal microbiota following antibiotic perturbation. Infection and Immunity 77(6), 2367-2375.
Arslan, D., Steinbusch, K.J.J., Diels, L., De Wever, H., Buisman, C.J.N. and Hamelers, H.V.M. (2012) Effect of hydrogen and carbon dioxide on carboxylic acids patterns in mixed culture fermentation. Bioresource Technology 118(0), 227-234.
Arslan, D., Steinbusch, K.J.J., Diels, L., De Wever, H., Hamelers, H.V.M. and Buisman, C.J.N. (2013) Selective carboxylate production by controlling hydrogen, carbon dioxide and substrate concentrations in mixed culture fermentation. Bioresource Technology 136(0), 452-460.
Arslan, D., Steinbusch, K.J.J., Diels, L., Hamelers, H.V.M., Strik, D.P.B.T.B., Buisman, C.J.N. and De Wever, H. (2016) Selective short chain carboxylates production: a review on control mechanisms to direct mixed culture fermentations. Critical Reviews in Environmental Science and Technology 46(6), 592-634.
Barberán, A., Bates, S.T., Casamayor, E.O. and Fierer, N. (2011) Using network analysis to explore co-occurrence patterns in soil microbial communities. The ISME journal 6(2), 343-351.
Barker, H. and Taha, S. (1942) *Clostridium kluyveri*, an organism concerned in the formation of caproic acid from ethyl alcohol. Journal of Bacteriology 43(3), 347.
Barker, H.A., Kamen, M.D. and Bornstein, B.T. (1945) The synthesis of *n*-butyric and *n*-caproic acids from ethanol and acetic acid by *Clostridium kluyveri*. Proceedings of the National Academy of Sciences 31(12), 373-381.
Bastidas-Oyanedel, J.-R., Mohd-Zaki, Z., Zeng, R.J., Bernet, N., Pratt, S., Steyer, J.-P. and Batstone, D.J. (2012) Gas controlled hydrogen fermentation. Bioresource Technology 110(0), 503-509.
Beckers, L., Masset, J., Hamilton, C., Delvigne, F., Toye, D., Crine, M., Thonart, P. and Hiligsmann, S. (2015) Investigation of the links between mass transfer conditions, dissolved hydrogen concentration and biohydrogen production by the pure strain *Clostridium butyricum* CWBI1009. Biochemical Engineering Journal 98, 18-28.
Biegel, E., Schmidt, S., González, J. and Müller, V. (2011) Biochemistry, evolution and physiological function of the Rnf complex, a novel ion-motive electron transport complex in prokaryotes. Cellular and Molecular Life Sciences 68(4), 613-634.
Bornstein, B.T. and Barker, H.A. (1948) The energy metabolism of *Clostridium kluyveri* and the synthesis of fatty acids. Journal of Biological Chemistry 172(2), 659-669.
Botton, S., van Heusden, M., Parsons, J.R., Smidt, H. and van Straalen, N. (2006) Resilience of microbial systems towards disturbances. Critical Reviews in Microbiology 32(2), 101-112.
Brink, G.-J.t., Arends, I.W.C.E. and Sheldon, R.A. (2000) Green, catalytic oxidation of alcohols in water. Science 287(5458), 1636-1639.
Buchfink, B., Xie, C. and Huson, D.H. (2015) Fast and sensitive protein alignment using DIAMOND. Nature Methods 12(1), 59-60.
Buckel, W. and Thauer, R.K. (2013) Energy conservation via electron bifurcating ferredoxin reduction and proton/Na^+ translocating ferredoxin oxidation. Biochimica et Biophysica Acta (BBA) - Bioenergetics 1827(2), 94-113.
Butkus, M.A., Hughes, K.T., Bowman, D.D., Liotta, J.L., Jenkins, M.B. and Labare, M.P. (2011) Inactivation of *Ascaris suum* by short-chain fatty acids. Applied and Environmental Microbiology 77(1), 363-366.
Butsch, B.M. and Bachofen, R. (1984) The membrane potential in whole cells of

Methanobacterium thermoautotrophicum. Archives of Microbiology 138(4), 293-298.
Callaway, T.R. and Russell, J.B. (1999) Selection of a highly monensin-resistant *Prevotella bryantii* subpopulation with altered outer membrane characteristics. Applied and Environmental Microbiology 65(11), 4753-4759.
Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., Fierer, N., Pena, A.G., Goodrich, J.K., Gordon, J.I., Huttley, G.A., Kelley, S.T., Knights, D., Koenig, J.E., Ley, R.E., Lozupone, C.A., McDonald, D., Muegge, B.D., Pirrung, M., Reeder, J., Sevinsky, J.R., Turnbaugh, P.J., Walters, W.A., Widmann, J., Yatsunenko, T., Zaneveld, J. and Knight, R. (2010) QIIME allows analysis of high-throughput community sequencing data. Nat Meth 7(5), 335-336.
Caporaso, J.G., Lauber, C.L., Walters, W.A., Berg-Lyons, D., Huntley, J., Fierer, N., Owens, S.M., Betley, J., Fraser, L., Bauer, M., Gormley, N., Gilbert, J.A., Smith, G. and Knight, R. (2012) Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. ISME J 6(8), 1621-1624.
Cavalcante, W.d.A., Leitão, R.C., Gehring, T.A., Angenent, L.T. and Santaella, S.T. (2017) Anaerobic fermentation for *n*-caproic acid production: A review. Process Biochemistry 54, 106-119.
Chen, M. and Wolin, M.J. (1979) Effect of monensin and lasaloid-sodium on the growth of methanogenic and saccharolytic bacteria. Applied Environmental Microbiology 38(1), 72-77.
Chen, Y., Cheng, J.J. and Creamer, K.S. (2008) Inhibition of anaerobic digestion process: A review. Bioresource Technology 99(10), 4044-4064.
Chow, C.-E.T., Sachdeva, R., Cram, J.A., Steele, J.A., Needham, D.M., Patel, A., Parada, A.E. and Fuhrman, J.A. (2013) Temporal variability and coherence of euphotic zone bacterial communities over a decade in the Southern California Bight. The ISME journal 7(12), 2259-2273.
Cline, M.S., Smoot, M., Cerami, E., Kuchinsky, A., Landys, N., Workman, C., Christmas, R., Avila-Campilo, I., Creech, M. and Gross, B. (2007) Integration of biological networks and gene expression data using Cytoscape. Nature protocols 2(10), 2366-2382.
Coleman, M.R., Moran, J.W. and Mowrey, D.H. (1997) Liquid chromatographic determination of monensin in premix and animal feeds: Collaborative study. Journal of AOAC International 80(4), 693-702.
CRC (2013) CRC Handbook of Chemistry and Physics, CRC Press.
Cuív, P.Ó., Smith, W.J., Pottenger, S., Burman, S., Shanahan, E.R. and Morrison, M. (2015) Isolation of genetically tractable most-wanted bacteria by metaparental mating. Scientific Reports 5, 13282.
Daniell, J., Köpke, M. and Simpson, S. (2012) Commercial biomass syngas fermentation. Energies 5(12), 5372.
De Kok, S., Meijer, J., Loosdrecht, M.M. and Kleerebezem, R. (2013) Impact of dissolved hydrogen partial pressure on mixed culture fermentations. Applied Microbiology and Biotechnology 97(6), 2617-2625.
de los Reyes Iii, F.L., Weaver, J.E. and Wang, L. (2015) A methodological framework for linking bioreactor function to microbial communities and environmental conditions. Current Opinion in Biotechnology 33, 112-118.
De Vrieze, J., Christiaens, M.E.R., Walraedt, D., Devooght, A., Ijaz, U.Z. and Boon, N. (2017) Microbial community redundancy in anaerobic digestion drives process recovery after

salinity exposure. Water Research 111, 109-117.
Dellomonaco, C., Clomburg, J.M., Miller, E.N. and Gonzalez, R. (2011) Engineered reversal of the beta-oxidation cycle for the synthesis of fuels and chemicals. Nature 476(7360), 355-359.
Demšar, J., Curk, T., Erjavec, A., Gorup, Č., Hočevar, T., Milutinovič, M., Možina, M., Polajnar, M., Toplak, M. and Starič, A. (2013) Orange: data mining toolbox in Python. The Journal of Machine Learning Research 14(1), 2349-2353.
Dennis, S.M. and Nagaraja, T.G. (1981) Effect of lasalocid or monensin on lactate-producing or using rumen bacteria. Journal of Animal Science 52(2), 418-426.
Desbois, A.P. (2012) Potential applications of antimicrobial fatty acids in medicine, agriculture and other industries. Recent patents on anti-infective drug discovery 7(2), 111-122.
Dethlefsen, L., Huse, S., Sogin, M.L. and Relman, D.A. (2008) The pervasive effects of an antibiotic on the human gut microbiota, as revealed by deep 16S rRNA sequencing. PLoS Biol 6(11), e280.
Diekert, G. and Wohlfarth, G. (1994) Metabolism of homoacetogens. Antonie van Leeuwenhoek 66(1-3), 209-221.
Diender, M., Stams, A.J.M. and Sousa, D.Z. (2016) Production of medium-chain fatty acids and higher alcohols by a synthetic co-culture grown on carbon monoxide or syngas. Biotechnology for Biofuels 9(1), 1-11.
Ding, H.-B., Tan, G.-Y.A. and Wang, J.-Y. (2010) *n*-Caproate formation in mixed-culture fermentative hydrogen production. Bioresource Technology 101(24), 9550-9559.
Dolliver, H., Gupta, S. and Noll, S. (2008) Antibiotic degradation during manure composting J. Environ. Qual. 37(3), 1245-1253.
Donoho, A., Manthey, J., Occolowitz, J. and Zornes, L. (1978) Metabolism of monensin in the steer and rat. Journal of Agricultural and Food Chemistry 26(5), 1090-1095.
Donoho, A.L. (1984) Biochemical studies on the fate of monensin in animals and in the environment. Journal of Animal Science 58(6), 1528-1539.
Eaton, A.D., Clesceri, L.S., Rice, E.W. and Greenberg, A.E. (eds) (2005) Standard methods for the examination of water and wastewater Washington, DC.
Edgar, R.C. (2010) Search and clustering orders of magnitude faster than BLAST. Bioinformatics 26(19), 2460-2461.
Fennell, D.E. and Gossett, J.M. (1998) Modeling the production of and competition for hydrogen in a dechlorinating culture. Environmental Science & Technology 32(16), 2450-2460.
Fernández, A., Huang, S., Seston, S., Xing, J., Hickey, R., Criddle, C. and Tiedje, J. (1999) How stable is stable? Function versus community composition. Applied and Environmental Microbiology 65(8), 3697-3704.
Fernandez, A.S., Hashsham, S.A., Dollhopf, S.L., Raskin, L., Glagoleva, O., Dazzo, F.B., Hickey, R.F., Criddle, C.S. and Tiedje, J.M. (2000) Flexible community structure correlates with stable community function in methanogenic bioreactor communities perturbed by glucose. Applied and Environmental Microbiology 66(9), 4058-4067.
Gaddy, J.L., Arora, D.K., Ko, C.W., Phillips, J.R., Basu, R., Wikstrom, C.V. and Clausen, E.C. (2007) Methods for increasing the production of ethanol from microbial fermentation, US7285402 B2, US Patents.
Gaertner, C.A., Serrano-Ruiz, J.C., Braden, D.J. and Dumesic, J.A. (2009) Catalytic coupling of carboxylic acids by ketonization as a processing step in biomass conversion. Journal of Catalysis 266(1), 71-78.

Ge, S., Usack, J.G., Spirito, C.M. and Angenent, L.T. (2015) Long-term *n*-caproic acid production from yeast-fermentation beer in an anaerobic bioreactor with continuous product extraction. Environ Sci Technol 49(13), 8012-8021.

Gilbert, J., Meyer, F., Antonopoulos, D.A., Balaji, P., Brown, C.T., Brown, C.T., Desai, N., Eisen, J.A., Evers, D., Field, D., Feng, W., Huson, D., Jansson, J., Knight, R., Knight, J., Kolker, E., Konstantindis, K., Kostka, J., Kyrpides, N.C., Mackelprang, R., McHardy, A., Quince, C., Raes, J., Sczyrba, A., Shade, A. and Stevens, R. (2010) Meeting report: The terabase metagenomics workshop and the vision of an earth microbiome project.

Gildemyn, S., Molitor, B., Usack, J.G., Nguyen, M., Rabaey, K. and Angenent, L.T. (2017) Upgrading syngas fermentation effluent using *Clostridium kluyveri* in a continuous fermentation. Biotechnology for Biofuels 10(1), 83.

Gonzalez-Cabaleiro, R., Lema, J.M., Rodriguez, J. and Kleerebezem, R. (2013) Linking thermodynamics and kinetics to assess pathway reversibility in anaerobic bioprocesses. Energy & Environmental Science 6(12), 3780-3789.

Gophna, U., Konikoff, T. and Nielsen, H.B. (2017) *Oscillospira* and related bacteria – From metagenomic species to metabolic features. Environmental Microbiology 19(3), 835-841.

Gottschalk, G. (1986) Bacterial metabolism, Springer-Verlag, New York.

Graef, S.P. and Andrews, J.F. (1974) Stability and control of anaerobic digestion. Journal (Water Pollution Control Federation) 46(4), 666-683.

Grootscholten, T.I.M., Kinsky dal Borgo, F., Hamelers, H.V.M. and Buisman, C.J.N. (2013a) Promoting chain elongation in mixed culture acidification reactors by addition of ethanol. Biomass and Bioenergy 48(0), 10-16.

Grootscholten, T.I.M., Steinbusch, K.J.J., Hamelers, H.V.M. and Buisman, C.J.N. (2013b) Chain elongation of acetate and ethanol in an upflow anaerobic filter for high rate MCFA production. Bioresource Technology 135(0), 440-445.

Grootscholten, T.I.M., Steinbusch, K.J.J., Hamelers, H.V.M. and Buisman, C.J.N. (2013c) High rate heptanoate production from propionate and ethanol using chain elongation. Bioresource Technology 136(0), 715-718.

Grootscholten, T.I.M., Steinbusch, K.J.J., Hamelers, H.V.M. and Buisman, C.J.N. (2013d) Improving medium chain fatty acid productivity using chain elongation by reducing the hydraulic retention time in an upflow anaerobic filter. Bioresource Technology 136(0), 735-738.

Grootscholten, T.I.M., Strik, D.P.B.T.B., Steinbusch, K.J.J., Buisman, C.J.N. and Hamelers, H.V.M. (2014) Two-stage medium chain fatty acid (MCFA) production from municipal solid waste and ethanol. Applied Energy 116(0), 223-229.

Hai, R., Wang, Y., Wang, X., Li, Y. and Du, Z. (2014) Bacterial community dynamics and taxa-time relationships within two activated sludge bioreactors. PLoS ONE 9(3), e90175.

Hanreich, A., Schimpf, U., Zakrzewski, M., Schlüter, A., Benndorf, D., Heyer, R., Rapp, E., Pühler, A., Reichl, U. and Klocke, M. (2013) Metagenome and metaproteome analyses of microbial communities in mesophilic biogas-producing anaerobic batch fermentations indicate concerted plant carbohydrate degradation. Systematic and Applied Microbiology 36(5), 330-338.

Hanselmann, K.W. (1991) Microbial energetics applied to waste repositories. Experientia 47(7), 645-687.

Hansson, G. and Molin, N. (1981) End product inhibition in methane fermentations: Effects of carbon dioxide on fermentative and acetogenic bacteria. European Journal of Applied

Microbiology and Biotechnology 13(4), 242-247.
Harper, S.R. and Pohland, F.G. (1986) Recent developments in hydrogen management during anaerobic biological wastewater treatment. Biotechnology and Bioengineering 28(4), 585-602.
Hartmann, M., Niklaus, P.A., Zimmermann, S., Schmutz, S., Kremer, J., Abarenkov, K., Luscher, P., Widmer, F. and Frey, B. (2014) Resistance and resilience of the forest soil microbiome to logging-associated compaction. The ISME journal 8(1), 226-244.
Harvey, B.G. and Meylemans, H.A. (2014) 1-Hexene: a renewable C6 platform for full-performance jet and diesel fuels. Green Chemistry 16(2), 770-776.
Hashsham, S.A., Fernandez, A.S., Dollhopf, S.L., Dazzo, F.B., Hickey, R.F., Tiedje, J.M. and Criddle, C.S. (2000) Parallel processing of substrate correlates with greater functional stability in methanogenic bioreactor communities perturbed by glucose. Applied and Environmental Microbiology 66(9), 4050-4057.
Hastie, T., Tibshirani, R., Narasimhan, B. and Chu, G. (2013) pamr: Pam: prediction analysis for microarrays. R package 1.
Hattori, S. (2008) Syntrophic acetate-oxidizing microbes in methanogenic environments. Microbes and Environments 23(2), 118-127.
Herrmann, G., Jayamani, E., Mai, G. and Buckel, W. (2008) Energy conservation *via* electron-transferring flavoprotein in anaerobic bacteria. Journal of Bacteriology 190(3), 784-791.
Hilpert, R., Winter, J. and Kandler, O. (1984) Agricultural feed additives and disinfectants as inhibitory factors in anaerobic digestion. Agricultural Wastes 10(2), 103-116.
Hollander, M. and Wolfe, D.A. (1973) Nonparametric Statistical Methods, Wiley, New York.
Hollister, E., Forrest, A., Wilkinson, H., Ebbole, D., Malfatti, S., Tringe, S., Holtzapple, M. and Gentry, T. (2010) Structure and dynamics of the microbial communities underlying the carboxylate platform for biofuel production. Applied Microbiology and Biotechnology 88(1), 389-399.
Hollister, E.B., Forrest, A.K., Wilkinson, H.H., Ebbole, D.J., Tringe, S.G., Malfatti, S.A., Holtzapple, M.T. and Gentry, T.J. (2012) Mesophilic and thermophilic conditions select for unique but highly parallel microbial communities to perform carboxylate platform biomass conversion. PLoS ONE 7(6), e39689.
Holtzapple, M., Davison, R., Ross, M., Aldrett-Lee, S., Nagwani, M., Lee, C.-M., Lee, C., Adelson, S., Kaar, W., Gaskin, D., Shirage, H., Chang, N.-S., Chang, V. and Loescher, M. (1999) Biomass conversion to mixed alcohol fuels using the MixAlco process. Applied Biochemistry and Biotechnology 79(1), 609-631.
Holtzapple, M. and Granda, C. (2009) Carboxylate platform: the MixAlco process part 1: comparison of three biomass conversion platforms. Applied Biochemistry and Biotechnology 156(1-3), 95-106.
Huang, H.J. and Ramaswamy, S. (2013) Overview of biomass conversion processes and separation and purification technologies in biorefineries. Separation and Purification Technologies in Biorefineries. John Wiley and Sons, Ltd, 1-36.
Huson, D.H., Mitra, S., Ruscheweyh, H.-J., Weber, N. and Schuster, S.C. (2011) Integrative analysis of environmental sequences using MEGAN4. Genome Research 21(9), 1552-1560.
Imachi, H., Sakai, S., Ohashi, A., Harada, H., Hanada, S., Kamagata, Y. and Sekiguchi, Y. (2007) *Pelotomaculum propionicicum* sp. nov., an anaerobic, mesophilic, obligately syntrophic, propionate-oxidizing bacterium. International Journal of Systematic and Evolutionary

Microbiology 57(7), 1487-1492.
Inanc, B., Matsui, S. and Ide, S. (1999) Propionic acid accumulation in anaerobic digestion of carbohydrates: An investigation on the role of hydrogen gas. Water Science and Technology 40(1), 93-100.
Jackson, B.E. and McInerney, M.J. (2002) Anaerobic microbial metabolism can proceed close to thermodynamic limits. Nature 415(6870), 454-456.
Jaenicke, S., Ander, C., Bekel, T., Bisdorf, R., Dröge, M., Gartemann, K.-H., Jünemann, S., Kaiser, O., Krause, L., Tille, F., Zakrzewski, M., Pühler, A., Schlüter, A. and Goesmann, A. (2011) Comparative and joint analysis of two metagenomic datasets from a biogas fermenter obtained by 454-pyrosequencing. PLoS ONE 6(1), e14519.
Jeon, B., Kim, B.-C., Um, Y. and Sang, B.-I. (2010) Production of hexanoic acid from d-galactitol by a newly isolated *Clostridium sp. BS-1*. Applied Microbiology and Biotechnology 88(5), 1161-1167.
Jeon, B.S., Choi, O., Um, Y. and Sang, B.-I. (2016) Production of medium-chain carboxylic acids by *Megasphaera* sp. MH with supplemental electron acceptors. Biotechnology for Biofuels 9(1), 129.
Jeon, B.S., Moon, C., Kim, B.-C., Kim, H., Um, Y. and Sang, B.-I. (2013) In situ extractive fermentation for the production of hexanoic acid from galactitol by *Clostridium sp. BS-1*. Enzyme and Microbial Technology 53(3), 143-151.
Johnson, M., Zaretskaya, I., Raytselis, Y., Merezhuk, Y., McGinnis, S. and Madden, T.L. (2008) NCBI BLAST: a better web interface. Nucleic Acids Research 36(suppl_2), W5-W9.
Jones, R.P. and Greenfield, P.F. (1982) Effect of carbon dioxide on yeast growth and fermentation. Enzyme and Microbial Technology 4(4), 210-223.
Jungermann, K., Thauer, R.K., Rupprecht, E., Ohrloff, C. and Decker, K. (1969) Ferredoxin mediated hydrogen formation from NADPH in a cell-free system of *Clostridium kluyveri*. FEBS Letters 3(2), 144-146.
Kampmann, K., Ratering, S., Kramer, I., Schmidt, M., Zerr, W. and Schnell, S. (2012) Unexpected stability of Bacteroidetes and Firmicutes communities in laboratory biogas reactors fed with different defined substrates. Applied and Environmental Microbiology 78(7), 2106-2119.
Kassen, R. (2002) The experimental evolution of specialists, generalists, and the maintenance of diversity. Journal of Evolutionary Biology 15(2), 173-190.
Katano, Y., Fujinami, S., Kawakoshi, A., Nakazawa, H., Oji, S., Iino, T., Oguchi, A., Ankai, A., Fukui, S., Terui, Y., Kamata, S., Harada, T., Tanikawa, S., Suzuki, K.-i. and Fujita, N. (2012) Complete genome sequence of *Oscillibacter valericigenes* Sjm18-20(T) (=NBRC 101213(T)). Standards in Genomic Sciences 6(3), 406-414.
Kenealy, W. and Waselefsky, D. (1985) Studies on the substrate range of *Clostridium kluyveri*; the use of propanol and succinate. Archives of Microbiology 141(3), 187-194.
Kenealy, W.R., Cao, Y. and Weimer, P.J. (1995) Production of caproic acid by cocultures of ruminal cellulolytic bacteria and *Clostridium kluyveri* grown on cellulose and ethanol. Applied Microbiology and Biotechnology 44(3), 507-513.
Kim, D.-H., Han, S.-K., Kim, S.-H. and Shin, H.-S. (2006) Effect of gas sparging on continuous fermentative hydrogen production. International Journal of Hydrogen Energy 31(15), 2158-2169.
Kim, M., Felix, T., Loerch, S. and Yu, Z. (2014) Effect of haylage and monensin supplementation on ruminal bacterial communities of feedlot cattle. Current Microbiology 69(2), 169-175.

Kleerebezem, R. and Stams, A.J.M. (2000) Kinetics of syntrophic cultures: A theoretical treatise on *n*-butyrate fermentation. Biotechnology and Bioengineering 67(5), 529-543.
Kleerebezem, R. and Van Loosdrecht, M.C.M. (2010) A generalized method for thermodynamic state analysis of environmental systems. Critical Reviews in Environmental Science and Technology 40(1), 1-54.
Koch, C., Müller, S., Harms, H. and Harnisch, F. (2014) Microbiomes in bioenergy production: From analysis to management. Current Opinion in Biotechnology 27(0), 65-72.
Kopylova, E., Noé, L. and Touzet, H. (2012) SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. Bioinformatics 28(24), 3211-3217.
Kovács, E., Wirth, R., Maróti, G., Bagi, Z., Rákhely, G. and Kovács, K.L. (2013) Biogas production from protein-rich biomass: Fed-batch anaerobic fermentation of casein and of pig blood and associated changes in microbial community composition. PLoS ONE 8(10), e77265.
Kraemer, J. and Bagley, D. (2006) Supersaturation of dissolved H_2 and CO_2 during fermentative hydrogen production with N_2 sparging. Biotechnology Letters 28(18), 1485-1491.
Kucek, L., Xu, J., Nguyen, M. and Angenent, L. (2016a) Waste conversion into *n*-caprylate and *n*-caproate: resource recovery from wine lees using anaerobic reactor microbiomes and in-line extraction. Frontiers in Microbiology 7, 1892.
Kucek, L.A., Nguyen, M. and Angenent, L.T. (2016b) Conversion of l-lactate into *n*-caproate by a continuously fed reactor microbiome. Water Research 93, 163-171.
Kucek, L.A., Spirito, C.M. and Angenent, L.T. (2016c) High *n*-caprylate productivities and specificities from dilute ethanol and acetate: chain elongation with microbiomes to upgrade products from syngas fermentation. Energy & Environmental Science 9(11), 3482-3494.
Langer, S.G., Ahmed, S., Einfalt, D., Bengelsdorf, F.R. and Kazda, M. (2015) Functionally redundant but dissimilar microbial communities within biogas reactors treating maize silage in co-fermentation with sugar beet silage. Microbial Biotechnology 8(5), 828-836.
Latif, H., Zeidan, A.A., Nielsen, A.T. and Zengler, K. (2014) Trash to treasure: Syngas fermentation for the production of biofuels and commodity chemicals. Current Opinion in Biotechnology 27, 79-87.
Leang, C., Ueki, T., Nevin, K.P. and Lovley, D.R. (2013) A genetic system for *Clostridium ljungdahlii*: a chassis for autotrophic production of biocommodities and a model homoacetogen. Applied and Environmental Microbiology 79(4), 1102-1109.
Levy, P.F., Sanderson, J.E. and Cheng, L.K. (1984) Kolbe electrolysis of mixtures of aliphatic organic acids. Journal of The Electrochemical Society 131(4), 773-777.
Levy, P.F., Sanderson, J.E., Kispert, R.G. and Wise, D.L. (1981) Biorefining of biomass to liquid fuels and organic chemicals. Enzyme and Microbial Technology 3(3), 207-215.
Li, D., Liu, C.-M., Luo, R., Sadakane, K. and Lam, T.-W. (2015a) MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. Bioinformatics 31(10), 1674-1676.
Li, F., Hinderberger, J., Seedorf, H., Zhang, J., Buckel, W. and Thauer, R.K. (2008) Coupled ferredoxin and crotonyl coenzyme A (CoA) reduction with NADH catalyzed by the butyryl-CoA dehydrogenase/Etf complex from *Clostridium kluyveri*. Journal of Bacteriology 190(3), 843-850.
Li, J., Rui, J., Yao, M., Zhang, S., Yan, X., Wang, Y., Yan, Z. and Li, X. (2015b) Substrate type and free ammonia determine bacterial community structure in full-scale mesophilic anaerobic digesters treating cattle or swine manure. Frontiers in Microbiology 6, 1337.

Liew, F., Martin, M.E., Tappel, R.C., Heijstra, B.D., Mihalcea, C. and Köpke, M. (2016) Gas fermentation-a flexible platform for commercial scale production of low-carbon-fuels and chemicals from waste and renewable feedstocks. Front Microbiol. 7.

Liu, Y., Lü, F., Shao, L. and He, P. (2016) Alcohol-to-acid ratio and substrate concentration affect product structure in chain elongation reactions initiated by unacclimatized inoculum. Bioresource Technology 218, 1140-1150.

Lonkar, S., Fu, Z. and Holtzapple, M. (2016) Optimum alcohol concentration for chain elongation in mixed-culture fermentation of cellulosic substrate. Biotechnology and Bioengineering 113(12), 2597-2604.

Lopetuso, L.R., Scaldaferri, F., Petito, V. and Gasbarrini, A. (2013) Commensal Clostridia: leading players in the maintenance of gut homeostasis. Gut Pathogens 5(1), 23.

Lozupone, C., Hamady, M. and Knight, R. (2006) UniFrac--an online tool for comparing microbial community diversity in a phylogenetic context. BMC Bioinformatics 7, 371.

Lozupone, C., Lladser, M.E., Knights, D., Stombaugh, J. and Knight, R. (2010) UniFrac: an effective distance metric for microbial community comparison. The ISME journal 5(2), 169-172.

MacLeod, F.A., Guiot, S.R. and Costerton, J.W. (1990) Layered structure of bacterial aggregates produced in an upflow anaerobic sludge bed and filter reactor. Applied and Environmental Microbiology 56(6), 1598-1607.

Marshall, C.W., LaBelle, E.V. and May, H.D. (2013) Production of fuels and chemicals from waste by microbiomes. Current Opinion in Biotechnology 24(3), 391-397.

Mavrovouniotis, M.L. (1990) Group contributions for estimating standard gibbs energies of formation of biochemical compounds in aqueous solution. Biotechnology and Bioengineering 36(10), 1070-1082.

McDonald, D., Price, M.N., Goodrich, J., Nawrocki, E.P., DeSantis, T.Z., Probst, A., Andersen, G.L., Knight, R. and Hugenholtz, P. (2012) An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. The ISME journal 6(3), 610-618.

McInerney, M. and Wofford, N. (1992) Enzymes involved in crotonate metabolism in *Syntrophomonas wolfei*. Archives of Microbiology 158(5), 344-349.

McInerney, M.J., Rohlin, L., Mouttaki, H., Kim, U., Krupp, R.S., Rios-Hernandez, L., Sieber, J., Struchtemeyer, C.G., Bhattacharyya, A., Campbell, J.W. and Gunsalus, R.P. (2007) The genome of *Syntrophus aciditrophicus*: Life at the thermodynamic limit of microbial growth. Proceedings of the National Academy of Sciences 104(18), 7600-7605.

Meyer, F., Paarmann, D., D'Souza, M., Olson, R., Glass, E., Kubal, M., Paczian, T., Rodriguez, A., Stevens, R., Wilke, A., Wilkening, J. and Edwards, R. (2008) The metagenomics RAST server - a public resource for the automatic phylogenetic and functional analysis of metagenomes. BMC Bioinformatics 9(1), 386.

Miller, J.H., Novak, J.T., Knocke, W.R. and Pruden, A. (2016) Survival of antibiotic resistant bacteria and horizontal gene transfer control antibiotic resistance gene content in anaerobic digesters. Frontiers in Microbiology 7, 263.

Molitor, B., Richter, H., Martin, M.E., Jensen, R.O., Juminaga, A., Mihalcea, C. and Angenent, L.T. (2016) Carbon recovery by fermentation of CO-rich off gases – Turning steel mills into biorefineries. Bioresource Technology 215, 386-396.

Morehead, M.C. and Dawson, K.A. (1992) Some growth and metabolic characteristics of monensin-sensitive and monensin-resistant strains of *Prevotella* (Bacteroides) *ruminicola*.

Applied and Environmental Microbiology 58(5), 1617-1623.
Munasinghe, P.C. and Khanal, S.K. (2010) Biomass-derived syngas fermentation into biofuels: Opportunities and challenges. Bioresource Technology 101(13), 5013-5022.
Mustapha, N.A., Sakai, K., Shirai, Y. and Maeda, T. (2016) Impact of different antibiotics on methane production using waste-activated sludge: Mechanisms and microbial community dynamics. Applied Microbiology and Biotechnology 100(21), 9355-9364.
Narihiro, T., Nobu, M.K., Kim, N.-K., Kamagata, Y. and Liu, W.-T. (2015) The nexus of syntrophy-associated microbiota in anaerobic digestion revealed by long-term enrichment and community survey. Environmental Microbiology 17(5), 1707-1720.
Narihiro, T., Terada, T., Ohashi, A., Kamagata, Y., Nakamura, K. and Sekiguchi, Y. (2012) Quantitative detection of previously characterized syntrophic bacteria in anaerobic wastewater treatment systems by sequence-specific rRNA cleavage method. Water Research 46(7), 2167-2175.
Nelson, M.C., Morrison, M. and Yu, Z. (2011) A meta-analysis of the microbial diversity observed in anaerobic digesters. Bioresource Technology 102(4), 3730-3739.
Oksanen, J., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O'Hara, R.B., Simpson, G.L., Solymos, P., Stevens, H.H. and Wagner, H. (2015) Vegan: Community ecology package. R package version 2.3-0.
Palmqvist, E. and Hahn-Hägerdal, B. (2000) Fermentation of lignocellulosic hydrolysates. II: inhibitors and mechanisms of inhibition. Bioresource Technology 74(1), 25-33.
Pauss, A., Andre, G., Perrier, M. and Guiot, S.R. (1990) Liquid-to-gas mass transfer in anaerobic processes: inevitable transfer limitations of methane and hydrogen in the biomethanation process. Applied and Environmental Microbiology 56(6), 1636-1644.
Pelletier, E., Kreimeyer, A., Bocs, S., Rouy, Z., Gyapay, G., Chouari, R., Rivière, D., Ganesan, A., Daegelen, P., Sghir, A., Cohen, G.N., Médigue, C., Weissenbach, J. and Le Paslier, D. (2008) "*Candidatus Cloacamonas Acidaminovorans*": Genome sequence reconstruction provides a first glimpse of a new bacterial division. Journal of Bacteriology 190(7), 2572-2579.
Phillips, J.R., Klasson, K.T., Clausen, E.C. and Gaddy, J.L. (1993) Biological production of ethanol from coal synthesis gas. Applied Biochemistry and Biotechnology 39(1), 559-571.
Poehlein, A., Schmidt, S., Kaster, A.-K., Goenrich, M., Vollmers, J., Thürmer, A., Bertsch, J., Schuchmann, K., Voigt, B., Hecker, M., Daniel, R., Thauer, R.K., Gottschalk, G. and Müller, V. (2012) An ancient pathway combining carbon dioxide fixation with the generation and utilization of a sodium ion gradient for ATP synthesis. PLoS ONE 7(3), e33439.
Press, W.H., Flannery, B.P., Teukolsky, S.A. and Vetterling, W.T. (1988) Numerical Recipes in C, Cambridge University Press, Cambridge.
R Core Team (2016) R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria.
Ragsdale, S.W. (2008) Enzymology of the Wood–Ljungdahl pathway of acetogenesis. Annals of the New York Academy of Sciences 1125(1), 129-136.
Regueiro, L., Spirito, C.M., Usack, J.G., Hospodsky, D., Werner, J.J. and Angenent, L.T. (2015) Comparing the inhibitory thresholds of dairy manure co-digesters after prolonged acclimation periods: Part 2 – correlations between microbiomes and environment. Water Research 87, 458-466.
Renz, M. (2005) Ketonization of carboxylic acids by decarboxylation: mechanism and scope.

European Journal of Organic Chemistry 2005(6), 979-988.
Richter, H., Martin, M.E. and Angenent, L.T. (2013) A two-stage continuous fermentation system for conversion of syngas into ethanol. Energies 6(8), 3987-4000.
Richter, H., Molitor, B., Diender, M., Sousa, D.Z. and Angenent, L.T. (2016) A narrow pH range supports butanol, hexanol, and octanol production from syngas in a continuous co-culture of *Clostridium ljungdahlii* and *Clostridium kluyveri* with in-line product extraction. Frontiers in Microbiology 7, 1773.
Riviere, D., Desvignes, V., Pelletier, E., Chaussonnerie, S., Guermazi, S., Weissenbach, J., Li, T., Camacho, P. and Sghir, A. (2009) Towards the definition of a core of microorganisms involved in anaerobic digestion of sludge. The ISME journal 3(6), 700-714.
Rodríguez, J., Lema, J.M. and Kleerebezem, R. (2008) Energy-based models for environmental biotechnology. Trends in Biotechnology 26(7), 366-374.
RStudio Team (2016) RStudio: Integrated development environment for R, RStudio, Inc., Boston, MA.
Russell, J.B. and Martin, S.A. (1984) Effects of various methane inhibitors on the fermentation of amino acids by mixed rumen microorganisms in vitro. Journal of Animal Science 59(5), 1329-1338.
Rychlik, J.L. and Russell, J.B. (2002) The adaptation and resistance of *Clostridium aminophilum* F to the butyrivibriocin-like substance of *Butyrivibrio fibrisolvens* JL5 and monensin. FEMS Microbiology Letters 209(1), 93-98.
Sangwan, N., Xia, F. and Gilbert, J.A. (2016) Recovering complete and draft population genomes from metagenome datasets. Microbiome 4(1), 8.
Sassman, S.A. and Lee, L.S. (2007) Sorption and degradation in soils of veterinary ionophore antibiotics: monensin and lasalocid. Environ Toxicol Chem 26(8), 1614-1621.
Schiel-Bengelsdorf, B. and Dürre, P. (2012) Pathway engineering and synthetic biology using acetogens. FEBS Letters 586(15), 2191-2198.
Schink, B. (1997) Energetics of syntrophic cooperation in methanogenic degradation. Microbiology and Molecular Biology Reviews 61(2), 262-280.
Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B., Lesniewski, R.A., Oakley, B.B., Parks, D.H., Robinson, C.J., Sahl, J.W., Stres, B., Thallinger, G.G., Van Horn, D.J. and Weber, C.F. (2009) Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. Applied and Environmental Microbiology 75(23), 7537-7541.
Seedorf, H., Fricke, W.F., Veith, B., Brüggemann, H., Liesegang, H., Strittmatter, A., Miethke, M., Buckel, W., Hinderberger, J., Li, F., Hagemeier, C., Thauer, R.K. and Gottschalk, G. (2008) The genome of *Clostridium kluyveri*, a strict anaerobe with unique metabolic features. Proceedings of the National Academy of Sciences 105(6), 2128-2133.
Shade, A., Caporaso, G.J., Handelsman, J., Knight, R. and Fierer, N. (2013) A meta-analysis of changes in bacterial and archaeal communities with time. The ISME journal 7(8), 1493-1506.
Shade, A., Peter, H., Allison, S.D., Baho, D., Berga, M., Buergmann, H., Huber, D.H., Langenheder, S., Lennon, J.T., Martiny, J.B., Matulich, K.L., Schmidt, T.M. and Handelsman, J. (2012) Fundamentals of microbial community resistance and resilience. Frontiers in Microbiology 3, 417
Shade, A., Read, J.S., Welkie, D.G., Kratz, T.K., Wu, C.H. and McMahon, K.D. (2011) Resistance, resilience and recovery: aquatic bacterial dynamics after water column

disturbance. Environmental Microbiology 13(10), 2752-2767.
Shannon, C.E. (1948) A mathematical theory of communication. Bell Syst. Tech. J. 27, 379-423.
Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B. and Ideker, T. (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Research 13(11), 2498-2504.
Simjee, S., Heffron, A.-L., Pridmore, A. and Shryock, T.R. (2012) Reversible monensin adaptation in *Enterococcus faecium*, *Enterococcus faecalis* and *Clostridium perfringens* of cattle origin: potential impact on human food safety. Journal of Antimicrobial Chemotherapy 67(10), 2388-2395.
Skřivanová, E. and Marounek, M. (2007) Influence of pH on antimicrobial activity of organic acids against rabbit enteropathogenic strain of *Escherichia coli*. Folia Microbiologica 52(1), 70-72.
Smith, D.P. and McCarty, P.L. (1989a) Energetic and rate effects on methanogenesis of ethanol and propionate in perturbed CSTRs. Biotechnology and Bioengineering 34(1), 39-54.
Smith, D.P. and McCarty, P.L. (1989b) Reduced product formation following perturbation of ethanol- and propionate-fed methanogenic CSTRs. Biotechnology and Bioengineering 34(7), 885-895.
Smith, G.M., Kim, B.W., Franke, A.A. and Roberts, J.D. (1985) 13C NMR studies of *n*-butyric fermentation in *Clostridium kluyveri*. Journal of Biological Chemistry 260(25), 13509-13512.
Speece, R.E. (1996) Anaerobic Biotechnology for Industrial Wastewaters, Archae Press.
Spirito, C.M., Richter, H., Rabaey, K., Stams, A.J.M. and Angenent, L.T. (2014) Chain elongation in anaerobic reactor microbiomes to recover resources from waste. Current Opinion in Biotechnology 27(0), 115-122.
Stams, A.J. and Plugge, C.M. (2009) Electron transfer in syntrophic communities of anaerobic bacteria and archaea. Nature Reviews Microbiology 7(8), 568-577.
Stein, R.R., Bucci, V., Toussaint, N.C., Buffie, C.G., Rätsch, G., Pamer, E.G., Sander, C. and Xavier, J.B. (2013) Ecological modeling from time-series inference: insight into dynamics and stability of intestinal microbiota. PLoS Comput Biol 9(12), e1003388.
Steinbusch, K.J., Hamelers, H.V.M., Plugge, C.M. and Buisman, C.J.N. (2011) Biological formation of *n*-caproate and *n*-caprylate from acetate: fuel and chemicals from low grade biomass. Energy & Environmental Science 4(1), 216-224.
Sträuber, H., Schröder, M. and Kleinsteuber, S. (2012) Metabolic and microbial community dynamics during the hydrolytic and acidogenic fermentation in a leach-bed process. Energy, Sustainability and Society 2(1), 1-10.
Sun, P., Huang, C.-H. and Pavlostathis, S.G. (2014) Inhibition and biotransformation potential of veterinary ionophore antibiotics under different redox conditions. Environmental Science & Technology 48(22), 13146-13154.
Sundberg, C., Al-Soud, W.A., Larsson, M., Alm, E., Yekta, S.S., Svensson, B.H., Sørensen, S.J. and Karlsson, A. (2013) 454 pyrosequencing analyses of bacterial and archaeal richness in 21 full-scale biogas digesters. FEMS Microbiology Ecology 85(3), 612-626.
Tao, Y., Li, J., Rui, J., Xu, Z., Zhou, Y., Hu, X., Wang, X., Liu, M., Li, D. and Li, X. (2014) Prokaryotic communities in pit mud from different-aged cellars used for the production of Chinese strong-flavored liquor. Applied and Environmental Microbiology 80(7), 2254-2260.
Teella, A., Huber, G.W. and Ford, D.M. (2011) Separation of acetic acid from the aqueous fraction

of fast pyrolysis bio-oils using nanofiltration and reverse osmosis membranes. Journal of Membrane Science 378(1–2), 495-502.
Thauer, R.K., Jungermann, K., Henninger, H., Wenning, J. and Decker, K. (1968) The energy metabolism of *Clostridium kluyveri*. European Journal of Biochemistry 4(2), 173-180.
Thauer, R.K., Jungermann, K., Rupprecht, E. and Decker, K. (1969) Hydrogen formation from NADH in cell-free extracts of *Clostridium kluyveri*: Acetyl coenzyme A requirement and ferredoxin dependence. FEBS Letters 4(2), 108-112.
Thaveesri, J., Boucneau, G. and Verstraete, W. (1994) Effect of monensin on UASB-reactor performance. Environmental Technology 15(5), 491-496.
Thomas, F., Hehemann, J.-H., Rebuffet, E., Czjzek, M. and Michel, G. (2011) Environmental and gut Bacteroidetes: The food connection. Frontiers in Microbiology 2, 93.
Thomas, T., Gilbert, J. and Meyer, F. (2012) Metagenomics-a guide from sampling to data analysis. Microbial Informatics and Experimentation 2(3).
Tomlinson, N. and Barker, H.A. (1954) Carbon dioxide and acetate utilization by *Clostridium kluyveri*. Journal of Biological Chemistry 209(2), 585-595.
Town, J.R., Links, M.G., Fonstad, T.A. and Dumonceaux, T.J. (2014) Molecular characterization of anaerobic digester microbial communities identifies microorganisms that correlate to reactor performance. Bioresource Technology 151(0), 249-257.
Tracy, B.P., Jones, S.W., Fast, A.G., Indurthi, D.C. and Papoutsakis, E.T. (2012) Clostridia: the importance of their exceptional substrate and metabolite diversity for biofuel and biorefinery applications. Current Opinion in Biotechnology 23(3), 364-381.
Tremblay, P.-L., Zhang, T., Dar, S.A., Leang, C. and Lovley, D.R. (2013) The Rnf complex of *Clostridium ljungdahlii* is a proton-translocating ferredoxin:NAD^+ oxidoreductase essential for autotrophic growth. mBio 4(1).
Usack, J.G. and Angenent, L.T. (2015) Comparing the inhibitory thresholds of dairy manure co-digesters after prolonged acclimation periods: Part 1 – Performance and operating limits. Water Research 87, 446-457.
Van Der Gast, C.J., Ager, D. and Lilley, A.K. (2008) Temporal scaling of bacterial taxa is influenced by both stochastic and deterministic ecological factors. Environmental Microbiology 10(6), 1411-1418.
Vane, L.M. (2008) Separation technologies for the recovery and dehydration of alcohols from fermentation broths. Biofuels, Bioproducts and Biorefining 2(6), 553-588.
Vanwonterghem, I., Jensen, P.D., Dennis, P.G., Hugenholtz, P., Rabaey, K. and Tyson, G.W. (2014a) Deterministic processes guide long-term synchronised population dynamics in replicate anaerobic digesters. The ISME journal 8(10), 2015-2028.
Vanwonterghem, I., Jensen, P.D., Ho, D.P., Batstone, D.J. and Tyson, G.W. (2014b) Linking microbial community structure, interactions and function in anaerobic digesters using new molecular techniques. Current Opinion in Biotechnology 27(0), 55-64.
Varel, V.H. and Hashimoto, A.G. (1981) Effect of dietary monensin or chlortetracycline on methane production from cattle waste. Applied and Environmental Microbiology 41(1), 29-34.
Varel, V.H. and Hashimoto, A.G. (1982) Methane production by fermentor cultures acclimated to waste from cattle fed monensin, lasalocid, salinomycin, or avoparcin. Applied and Environmental Microbiology 44(6), 1415-1420.
Varel, V.H., Wells, J.E., Shelver, W.L., Rice, C.P., Armstrong, D.L. and Parker, D.B. (2012) Effect of anaerobic digestion temperature on odour, coliforms and chlortetracycline in

swine manure or monensin in cattle manure. Journal of Applied Microbiology 112(4), 705-715.

Vasudevan, D., Richter, H. and Angenent, L.T. (2014) Upgrading dilute ethanol from syngas fermentation to *n*-caproate with reactor microbiomes. Bioresource Technology 151(0), 378-382.

Vuono, D.C., Benecke, J., Henkel, J., Navidi, W.C., Cath, T.Y., Munakata-Marr, J., Spear, J.R. and Drewes, J.E. (2015) Disturbance and temporal partitioning of the activated sludge metacommunity. The ISME journal 9(2), 425-435.

Wallace, R.J., Chaudhary, L.C., Miyagawa, E., McKain, N. and Walker, N.D. (2004) Metabolic properties of *Eubacterium pyruvativorans*, a ruminal 'hyper-ammonia-producing' anaerobe with metabolic properties analogous to those of *Clostridium kluyveri*. Microbiology 150(9), 2921-2930.

Wallace, R.J., McKain, N., McEwan, N.R., Miyagawa, E., Chaudhary, L.C., King, T.P., Walker, N.D., Apajalahti, J.H.A. and Newbold, C.J. (2003) *Eubacterium pyruvativorans sp. nov.*, a novel non-saccharolytic anaerobe from the rumen that ferments pyruvate and amino acids, forms caproate and utilizes acetate and propionate. International Journal of Systematic and Evolutionary Microbiology 53(4), 965-970.

Wang, J., Shen, J., Wu, Y., Tu, C., Soininen, J., Stegen, J.C., He, J., Liu, X., Zhang, L. and Zhang, E. (2013) Phylogenetic beta diversity in bacterial assemblages across ecosystems: deterministic versus stochastic processes. The ISME journal 7, 1310-1321.

Wang, Q., Garrity, G.M., Tiedje, J.M. and Cole, J.R. (2007) Naive bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Applied and Environmental Microbiology 73(16), 5261-5267.

Wang, S., Huang, H., Moll, J. and Thauer, R.K. (2010) $NADP^+$ reduction with reduced ferredoxin and $NADP^+$ reduction with NADH are coupled via an electron-bifurcating enzyme complex in *Clostridium kluyveri*. Journal of Bacteriology 192(19), 5115-5123.

Warnes, G.R., Bolker, B., Bonebakker, L., Gentleman, R., Huber, W., Liaw, A., Lumley, T., Maechler, M., Magnusson, A., Moeller, S., Schwartz, M. and Venables, B. (2015) gplots: Various R programming tools for plotting data. R package version 2.17.0.

Watanabe, N., Harter, T.H. and Bergamaschi, B.A. (2008) Environmental occurrence and shallow ground water detection of the antibiotic monensin from dairy farms. Journal of Environmental Quality 37(5_Supplement), S-78-S-85.

Weimer, P. and Stevenson, D. (2012) Isolation, characterization, and quantification of *Clostridium kluyveri* from the bovine rumen. Applied Microbiology and Biotechnology 94(2), 461-466.

Weimer, P., Stevenson, D., Mertens, D. and Thomas, E. (2008) Effect of monensin feeding and withdrawal on populations of individual bacterial species in the rumen of lactating dairy cows fed high-starch rations. Applied Microbiology and Biotechnology 80(1), 135-145.

Weimer, P.J. and Kohn, R.A. (2016) Impacts of ruminal microorganisms on the production of fuels: how can we intercede from the outside? Applied Microbiology and Biotechnology, 1-10.

Weimer, P.J., Nerdahl, M. and Brandl, D.J. (2015) Production of medium-chain volatile fatty acids by mixed ruminal microorganisms is enhanced by ethanol in co-culture with *Clostridium kluyveri*. Bioresource Technology 175, 97-101.

Werner, J.J., Garcia, M.L., Perkins, S.D., Yarasheski, K.E., Smith, S.R. and Muegge, B.D. (2014) Microbial community dynamics and stability during an ammonia-induced shift to syntrophic acetate oxidation. Applied and Environmental Microbiology 80(11), 3375-

3383.
Werner, J.J., Knights, D., Garcia, M.L., Scalfone, N.B., Smith, S., Yarasheski, K., Cummings, T.A., Beers, A.R., Knight, R. and Angenent, L.T. (2011) Bacterial community structures are unique and resilient in full-scale bioenergy systems. Proceedings of the National Academy of Sciences 108(10), 4158-4163.
Wirth, R., Kovács, E., Maróti, G., Bagi, Z., Rákhely, G. and Kovács, K.L. (2012) Characterization of a biogas-producing microbial community by short-read next generation DNA sequencing. Biotechnol Biofuels 5(4).
Wittebolle, L., Marzorati, M., Clement, L., Balloi, A., Daffonchio, D., Heylen, K., De Vos, P., Verstraete, W. and Boon, N. (2009) Initial community evenness favours functionality under selective stress. Nature 458(7238), 623-626.
Wrighton, K.C., Castelle, C.J., Wilkins, M.J., Hug, L.A., Sharon, I., Thomas, B.C., Handley, K.M., Mullin, S.W., Nicora, C.D., Singh, A., Lipton, M.S., Long, P.E., Williams, K.H. and Banfield, J.F. (2014) Metabolic interdependencies between phylogenetically novel fermenters and respiratory organisms in an unconfined aquifer. The ISME journal 8(7), 1452-1463.
Wrighton, K.C., Thomas, B.C., Sharon, I., Miller, C.S., Castelle, C.J., VerBerkmoes, N.C., Wilkins, M.J., Hettich, R.L., Lipton, M.S., Williams, K.H., Long, P.E. and Banfield, J.F. (2012) Fermentation, hydrogen, and sulfur metabolism in multiple uncultivated bacterial phyla. Science 337(6102), 1661-1665.
Wu, Z. and Yang, S.-T. (2003) Extractive fermentation for butyric acid production from glucose by *Clostridium tyrobutyricum*. Biotechnology and Bioengineering 82(1), 93-102.
Yamada, T. and Sekiguchi, Y. (2009) Cultivation of uncultured Chloroflexi subphyla: Significance and ecophysiology of formerly uncultured Chloroflexi 'Subphylum I' with natural and biotechnological relevance. Microbes and Environments 24(3), 205-216.
Youssef, N.H., Blainey, P.C., Quake, S.R. and Elshahed, M.S. (2011) Partial genome assembly for a candidate division OP11 single cell from an anoxic spring (Zodletone Spring, Oklahoma). Applied and Environmental Microbiology 77(21), 7804-7814.
Yu, H.-Q. and Mu, Y. (2006) Biological hydrogen production in a UASB reactor with granules. II: Reactor performance in 3-year operation. Biotechnology and Bioengineering 94(5), 988-995.
Zaks, D.P.M., Winchester, N., Kucharik, C.J., Barford, C.C., Paltsev, S. and Reilly, J.M. (2011) Contribution of anaerobic digesters to emissions mitigation and electricity generation under U.S. climate policy. Environmental Science & Technology 45(16), 6735-6742.
Zerbino, D.R. and Birney, E. (2008) Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. Genome Research 18(5), 821-829.
Zhang, F., Ding, J., Zhang, Y., Chen, M., Ding, Z.-W., van Loosdrecht, M.C.M. and Zeng, R.J. (2013a) Fatty acids production from hydrogen and carbon dioxide by mixed culture in the membrane biofilm reactor. Water Research 47(16), 6122-6129.
Zhang, F., Zhang, Y., Chen, M., van Loosdrecht, M.C.M. and Zeng, R.J. (2013b) A modified metabolic model for mixed culture fermentation with energy conserving electron bifurcation reaction and metabolite transport energy. Biotechnology and Bioengineering 110(7), 1884-1894.
Zhang, W., Werner, J.J., Agler, M.T. and Angenent, L.T. (2014) Substrate type drives variation in reactor microbiomes of anaerobic digesters. Bioresource Technology 151, 397-401.
Zhang, X., Li, C., Wang, Y., Luo, J. and Xu, T. (2011) Recovery of acetic acid from simulated

acetaldehyde wastewaters: Bipolar membrane electrodialysis processes and membrane selection. Journal of Membrane Science 379(1–2), 184-190.
Zhu, W., Lomsadze, A. and Borodovsky, M. (2010) Ab initio gene identification in metagenomic sequences. Nucleic Acids Research 38(12), e132-e132.
Zhu, X., Tao, Y., Liang, C., Li, X., Wei, N., Zhang, W., Zhou, Y., Yang, Y. and Bo, T. (2015a) The synthesis of *n*-caproate from lactate: a new efficient process for medium-chain carboxylates production. Scientific Reports 5, 14360.
Zhu, X., Zhou, Y., Wang, Y., Wu, T., Li, X., Li, D. and Tao, Y. (2017) Production of high-concentration *n*-caproic acid from lactate through fermentation using a newly isolated *Ruminococcaceae* bacterium CPB6. Biotechnology for Biofuels 10(1), 102.
Zhu, X.Y., Tao, Y., Liang, C., Li, X.Z., Wei, N. and Zhang, W.J. (2015b) The synthesis of *n*-caproate from lactate: a new efficient process for medium-chain carboxylates production. Sci Rep-Uk.
Zitomer, D., Burns, R., Duran, M., and Vogal, D. (2007) Effect of sanitizers, rumensin, and temperature on anaerobic digester biomass. Transactions of the ASABE 50(5), 1807-1813.
Zoetemeyer, R.J., van den Heuvel, J.C. and Cohen, A. (1982) pH influence on acidogenic dissimilation of glucose in an anaerobic digestor. Water Research 16(3), 303-311.